

Tue Feb 11 11:27:24 1997

US-08-674-311-1-01.rge

Page 2

```

source
1..2269
/clone="18-11 CDNA"
/map="9p21"
/chromosome="9"
/organism="Homo sapiens"
/sex="male"
/cell_line="primary culture"
/cell_type="fibroblast"
/tissue_type="epidermis"
122..973
/gene="MTAP"
/codon_start=1
/product="methylthioadenosine phosphorylase"
/db_xref="PID:9847724"
/translation="MASGTTTAVKIGLIGGTGLDPELIEGRTEKYVDTPEGKPSDALILGKIKNVDCILARHGROHTIMPSKYNVQANIMALKEEGCTHIVITTAGSLREEIQPDIVIIDQFIDRTMRPOSFYDGSFCARGVCHIPMAPEPCPTRVLIETAKLGLRCHSKGTMTWITEGPRFSSRAESFMRTGADVIMNTVPEVVLAKAEGICIASIAMAFIDYDCWKEHEEAVSDRVLKLTKENANKAKSILLTTIPQIGSTEMSETLHNLNMAQF SVLLPRH"

```

Query Match	100.0%;	Score 2269;	DB 56;	Length 2269;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;		
Matches 2269;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Db	1	gaatccgcctccgcactgctcaactcccgccagtgaggttgcacagaccacgcctctgtg	60
QY	1	GAATTCGGCTCCGCACtGCTCACTCCCGCCAGTGAGGTGGCACAGCCACGCTCTGTG	60
Db	61	gtctgcttggttcccttaagtcgccgagcgtctgcgccactgcgaattccttcccgltgcaga	120
QY	61	GCTCGCTTGGTTCCCTTAGTCCCGAGCGCTCGCCCACTGCAGATTCCCTTCCCGTGCAGA	120
Db	121	catggcctctgcaaccaaccaaccgcgctgaagattgnaataattgltgaaacaggcct	180
QY	121	CATGGCCTCTGGCACCAcCACCGCCGTGAAGATTGGATAATTGGTGAACAGGCCT	180
Db	181	ggaatgatccagaatattagaaggaagaaactgaaaaatatgttgatactccatttggcaa	240
QY	181	GGATGATCCAGAAATTTAGAAGGAAGAACTGAAAAATATGTGATACTCATTTGGCAA	240
Db	241	ggcaatctgatgccttaattttggggaagataaaaaatgttgaattgcacccctctgcaag	300
QY	241	GCCATCTGATGCTTATTTTGGGGAAGATAAAAAATGTGATTGCATCCTCTTGCAAG	300
Db	301	gcataggaaggcagcacaccatcatgtgccttcaaaagtcaactaccagcgnaacatctgggc	360
QY	301	GCATGGAAGGCAGACACCATCATGCTTCAAAAGTCAACTACCAAGCGAACAATCTGGGC	360
Db	361	tttgaaggaagagggcgtgtacacatgtcatagtgaccacagcttgtgtcctctgaggya	420
QY	361	TTTGAAGGAAGAGGGCTGTACACATGTCTATTAGTACCAcAGCTTGTGGCTCTTGAGGGA	420
Db	421	ggagatccaagcccgcgagatatgtcatatttgatcagttcatattgacaggaaccaatagag	480
QY	421	GGAGATTCAgCCCCGGCGATATTGTCAATTATGATCAGTTTATTGACAGGACCACCTATGAG	480
Db	481	aacctcaatccttcatatgatgtagaagtcattcttgtgcagagaggagtgtgcataattccaat	540
QY	481	ACCTCAgTCTTCTATGATGGAAGTCAATTCTTGTGCCAGAGAGTGTGCCATATTCCAAT	540
Db	541	ggctgagccgttttgcaccccaaaacgagagaggttctatagagactgtctaaagaagctagag	600
QY	541	GGCTGAGCCGTTTtGCCCCAAAACGAGAGGTTCTTATAGAGACTGCTAAAGAGCTAGG	600
Db	601	actccggtgcacctccaaggggagacaatggtcacaactgagggagcctcgttttagctcccg	660
QY	601	ACTCCGGTGCACACTCAAAAGGGAGACAATGTGCACAATGAGGGAGCCTCGTTTtagCTCCG	660
Db	661	ggcagaagcttcatgttccgcacctggggggcgagatgttatcaacaatgaaccacagttcc	720

QY	661	GGCAGAAAGCTTCATGTTCCGCACCTGGGGGGCGAGTGTATCAACATGACCACAGTTCC	720
Db	721	agaagtggttctctgctaaggagcgtggaattgttaacgaagtaacgcatgagcagaga	780
QY	721	AGAGGTGTTCTTGCTAAGGAGCGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA	780
Db	781	ttatgactgctggaaggagcagcaggaagcagtttcgggtggaacgggtccttaagaacct	840
QY	781	TTATGACTGCTGGAAGGAGCAGCAGAGAACGAGTTTCGGTGGACCGGCTCTTAAGACCCCT	840
Db	841	gaaagaaaaacgctaataaagccaaaagcttactgctcaactacatacctcagataggctc	900
QY	841	GAAAGAAAAACGCTAATAAAGCCAAAAAGCTTACTGCTCACTACCATACCTCAGATAGGGTC	900
Db	901	cacagaatggtcagaanaacctccataacctgaagaataygccccagtttctgtttatt	960
QY	901	CACAGATGGTCAGAAACCCTCCATTAACCTGAAGAAATATGGCCCACTTTCTGTTTTATT	960
Db	961	accagaacattaaagtagcatgctgcgccagagagaagaagaacattctaattccaagtc	1022
QY	961	ACCAAGACATTAAAGTAGCATGGCTGCCAGAGAAAAAGAGACATTCTAATCCAGTCA	1022
Db	1021	tttgggaattccctgcttaacttgaaaaaataatgggaagaacatgcaagcttcaagccc	1082
QY	1021	TTTTGGGAATTCCTGCTTAACCTGAAAAAATAATGGGAAAAGACATGCAGCTTTCATGCC	1082
Db	1081	ttgcctatcaaaagagtatgttgaagaagaacagacattgtygtattagaagactcctg	1142
QY	1081	TTGCCTATCAAAAGAGTATGTTGTAGAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG	1142
Db	1141	aatgatttagacaacttcaaatacagaagaanaagcaaatgactagtaaacatgtggaa	1200
QY	1141	AATGATTTAGACAACCTTCAAAATACAGAAAGAAAGCAAAATGACTAGTAACATGTGGAA	1200
Db	1201	aaataattacattttaagggggaaaaaaaccccaactctcttcctccctataa	1262
QY	1201	AAATATTACATTTTAAAGGGGGAATAAAAAAACCCACCATTTCTTCCCCCTATTAA	1262
Db	1261	atcgcaacaataaagggtgagaggtaactctacttccatactgccaagaatgtga	1322
QY	1261	ATTGCAACAATAAAGGTGAGGGTAACTCTTAATTCTTCCATACCTGCCAAGAATGTGA	1322
Db	1321	ggaagaaatgaggaactcttgtrattatgtgaatgagcagtaaatgtgtacagtaattct	1382
QY	1321	GGAAGAAATGGGACTCTTGGTTATTATTATGACGACTGTAATTGGTACAGTATTCT	1382
Db	1381	ggaaggcaatttggtaaaatgcatacaaaagacttaaaaatacaggaagctccttggtyctg	1442
QY	1381	GGAGGGCAATTTGCTAAATGCATCAAAAAGACTTAAAAATACGGACGCTCTTGGTGCTG	1442
Db	1441	ggaactctacatctagcaatttctctttaaaccatacagagatgcatacaaaagaatta	1502
QY	1441	GGAACCTACATCTAGCAATTTCTTTAAACCATATCAGAGATGCATACAAAGAAATTA	1502
Db	1501	tataaagaagggtgtttaataatgatatataataaataaattgaaacaactgta	1562
QY	1501	TATATAAAGAAAGGTGTTTAATAATGATAGTTAATAATAATTAATGAAACAATCTGA	1562
Db	1561	atcccttgcaatttggaggtaaatlatyctctagttataaactagattgtgaacagcaa	1622
QY	1561	ATCCCTGCAATTGGAGGTAAATATATGCTTAGTATATAATCTAGATTGTGAATCAGCCAA	1622
Db	1621	ctgaaaaatcccttttgcataattccaatgctcctaaagaacacgggtgctctatatatga	1682
QY	1621	CTGAATAATCCTTTTGCATATTTCATATGCTCTAAAAAGACACGGTGTCTTATATATGAA	1682
Db	1681	gtgaaaaaagatatgttagcaatttatagtaactagtcttgctttaaaatgctatgttaa	1742
QY	1681	GTGAAAAAAGATATGTAGCATTTTATAGTACTAGTTTGCCTTAAAAATGCTATGTAA	1742
Db	1741	tatacaaaaaaactagaagaataatataataaaccctgttaattgttattggggagagata	1802
QY	1741	TATACAAAAAACTAGAAAGAAATATATATAACCTTGTTATTTGTATTTGGGGAGGGATA	1802

QY	661	GGCAGAAAGCTTCATGTTCCGCACCTGGGGGGCGAGTGTATCAACATGACCACAGTTCC	720
Db	721	agaagtggttctctgctaaggagcgtggaattgttaacgaagtaacgcatgagcagaga	780
QY	721	AGAGGTGTTCTTGCTAAGGAGCGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA	780
Db	781	ttatgactgctggaaggagcagcaggaagcagtttcgggtggaacgggtccttaagaacct	840
QY	781	TTATGACTGCTGGAAGGAGCAGCAGAGAACGAGTTTCGGTGGACCGGCTCTTAAGACCCCT	840
Db	841	gaaagaaaaacgctaataaagccaaaagcttactgctcaactacatacctcagataggctc	900
QY	841	GAAAGAAAAACGCTAATAAAGCCAAAAAGCTTACTGCTCACTACCATCTCAGATAGGGTC	900
Db	901	cacagaatggtcagaanaacctccataacctgaagaataygccccagtttctgtttatt	960
QY	901	CACAGATGCTCAGAAACCCCTCCATTAACCTGAAGAAATATGGCCCACTTTCTGTTTTATT	960
Db	961	accagaacattaaagtagcatgctgcgccagagagaagaagaacattctaattccaagtc	1022
QY	961	ACCAAGACATTAAAGTAGCATGGCTGCCAGAGAAAAAGAGACATTCTAATCCAGTCA	1022
Db	1021	tttgggaattccctgcttaacttgaaaaaataatgggaagaacatgcaagcttcaagccc	1082
QY	1021	TTTTGGGAATTCCTGCTTAACCTGAAAAAATAATGGGAAAAGACATGCAGCTTTCATGCC	1082
Db	1081	ttgcctatcaaaagagtatgttgaagaagaacagacattgtygtattagaagactcctg	1142
QY	1081	TTGCCTATCAAAAGAGTATGTTGTAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG	1142
Db	1141	aatgatttagacaacttcaaatacagaagaanaagcaaatgactagtaaacatgtggaa	1200
QY	1141	AATGATTTAGACAACCTCAAAAATACAGAAAGAAAGCAAAATGACTAGTAACATGTGGAA	1200
Db	1201	aaataattacattttaaggagggaaaaaaaccccaactctcttcctccctataa	1262
QY	1201	AAATATTACATTTTAAAGGGGGAATAAAAAAACCCACCATGCTTCTCCCCCTATTAA	1262
Db	1261	atcgcaacaataaagggtgagaggtaactctcaacttccatactgccaagaatgtga	1322
QY	1261	ATTGCAACAATAAAGGTGAGGGTAATCTCTACTTCCATACTGCCAAAGAATGTGA	1322
Db	1321	ggaagaaatgaggaactcttggttattatgtgaatgagcagtaaatgtgtacagtaattct	1382
QY	1321	GGAAGAAATGGGACTCTTGGTTATTATTGATGCGACTGTAATTGGTACAGTATTCT	1382
Db	1381	ggaaggcaatttggtaaaatgcatacaaaagacttaaaaaatacaggaagctccttggtyctg	1442
QY	1381	GGAGGGCAATTTGCTAAAATGCATCAAAAAGACTTAAAAATACGGACGTCTTTGGTGTG	1442
Db	1441	ggaactctacatctagcaatttctctttaaaccatacagagatgcatacaaaagaatta	1502
QY	1441	GGAACCTACATCTAGCAATTTCTTTAAACCATATCAGAGATGCATACAAGAATTA	1502
Db	1501	tataaagaagggtgtttaataatgatatataataaataaattgaaacaactctga	1562
QY	1501	TATATAAGAAAGGTGTTTAATAATGATAGTATATAATAATAATTGAACAATCTGA	1562
Db	1561	atcccttgcaatttggaggtaaatlatyctctagttataaactagattgtgaacagcaa	1622
QY	1561	ATCCCTGCAATTGGAGGTAATATATGCTTAGTATATAATCTAGATTGTGAATCAGCCAA	1622
Db	1621	ctgaaaaatccttttgcataattccaatgctcctaaagaacacgggtgctctatatatga	1682
QY	1621	CTGAATAATCCTTTTGCATATTTCATATGCTCTAAAAAGACACGGTGTCTTATATATGAA	1682
Db	1681	gtgaaaaaagatatgttagcaatttatagtaactagtttggctttaaaatgctatgtaaa	1742
QY	1681	GTGAATAAGATATGTAGCATTTTATAGTACTAGTTTGCCTTAAAAATGCTATGTAAA	1742
Db	1741	tatacaaaaaactagaaagaatatatacaacctgttaattgttattgtggggaggagata	1802
QY	1741	TATACAAAAAACTAGAAAGAAATATATAAACCCTGTATATGTATTTGGGGAGGGATA	1802

Db 1801 ctgggataatttatacttcttgaatcttcctgctcctcacatttctacagtgaat 1860
|||||
OY 1801 CTGGGATAATTATTATTCTTTGGAATCTTCTGCTCTTACACATTTTCTACAGTGAAT 1860
Db 1861 ataacaataagtaagggccgtaaaaaaagtgtgattagaagaatccagttctga 1920
|||||
OY 1861 ATAATCAATAGTAAGGCCCGTAAAAAATAAGTGATTAGAAAGATCCAGTCTTGA 1920
Db 1921 aaacactgttcttgtaatgaagcagaatttaagttgtgtaataattaaggtgaatgtcatt 1980
|||||
OY 1921 AAACACTGTTCTGTGTAATGAAGCAGATTTAAGTTGTAATATAAGTGAATGTCAAT 1980
Db 1981 taaggagttacatcttattctctgctaaagaagagatcattgattctctgtacagtcaga 2040
|||||
OY 1981 TAAGGAGTTACATCTTTATTCTGCTAAGAGAAGATCATTGATTCTGTACAGTCAGA 2040
Db 2041 acagtaacttggtgtgcaacagcttctgagaaaaagctagtgataataagtttaactga 2100
|||||
OY 2041 ACAGTACTTGGGTGTCACACAGCTTTCGTGAGAAAAGCTAGGTGTATATAGTTTAACCTGA 2100
Db 2101 aagtttaactatttaaaagactaaatgacacatttatgtatctgatatatttaaaagta 2160
|||||
OY 2101 AAGTTTAACCTATTAAAGACTAAATAGCACATTTTATGTATCTGATATTTTAAAAAGTA 2160
Db 2161 atgtgagcttcctcttttatgagttaaatatttatcacgagttgtgtaattgtgcctt 2220
|||||
OY 2161 ATGTGAGCTTCCTCTTTTATGAGTTAAATATTATTATACGAGTGTGTAATTGTGCCTT 2220
Db 2221 ttaataaagtgtgaagctgtcttttataaaaaaagcggaattc 2269
|||||
OY 2221 TTAATAAGTGTGAAGCTTCTTTTAAAAAAGCGGAATTC 2269

RESULT 2
LOCUS HSPROKINX 3018 bp RNA PRI 01-JUN-1995
DEFINITION H.sapiens mRNA for Ndr protein kinase.
ACCESSION 235102
NID 9854169
KEYWORDS protein kinase.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Coelomata; Deuterostomia; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Homnidae; Homo.
1 (bases 1 to 3018)
REFERENCE
AUTHORS Millward,T.
TITLE Direct Submission
JOURNAL Submitted (06-JUL-1994) to the EMBL/GenBank/DBJ databases. Thomas
Millward, Friedrich Miescher-Institut, Basel, Postfach, 2543,
Switzerland
2 (bases 1 to 3018)
REFERENCE
AUTHORS Millward,T., Cron,P. and Hemmings,B.A.
TITLE Molecular cloning and characterization of a conserved nuclear
serine(threonine) protein kinase
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 92 (11), 5022-5026 (1995)
MEDLINE 95281588
FEATURES
source
Location/Qualifiers
1..3018
/organism="Homo sapiens"
/clone="human Ndr protein kinase"
/clone_1lb="Human fetal brain/Stratagene 936206"
596..1993
/codon_start=1
/product="Ndr protein kinase"
/db_xref="PID:9854170"
/translation="MAMTGSPPCSSMSNHTKERVTMTKVTLENFYSLIAOHEEREMR
OKLEKVEEGLKDEEKRLRSARHAKETFLRKTRRLGLEDFESLKYIGGAFGE
VRLVOKDGTGHVYAMKILRKADMLEKEQVGHIRAEKDILVEADSLWVVKMFYSFQDKL
NLYLIMEFLPGGDMTLLMKRDTLTEETQFYIAETVLAIDSIHQLGFHDIKPDNL
LDSKGVHKLSDFLCTGLKKAHRTETFYRNLNHSLPDSFTFQNMNSKRAETWKRNR

OLAFTVGTDPYIAPEVFMQTYNKLCDMWSLGVIMYEMLIGYPPRCSETPQETKKY
MMWKEUFTFPPEVPISEKADLILRFCEWEHRIGAPVEIKSSEFEGVDWEHIRE
RPAISIEIKSIDTSNDFEPESDILKPTVATSNHPETDYKKNQWVFINTYKRFEG
LTARCAIPSYMKAAR"
BASE COUNT 908 a 640 c 672 g 798 t
ORIGIN
Query Match 1.5%; Score 35; DB 55; Length 3018;
Best Local Similarity 74.0%; Pred. No. 2.62e-05;
Matches 54; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
Db 449 ttctctgcttcttggttacatatattacctgttctcctccagaaaaactgacccaatttac 508
|||||
CP 1419 TTTTAACTCTTTTGAATGCATTTTACCAAAATGCCCCCAGAAATACTGTACCAATTTAC 1360
Db 509 attcctaccaata 521
|||||
CP 1359 AGTCGCATCAATA 1347

RESULT 3
LOCUS SPAC16C9 14071 bp DNA PLN 22-OCT-1995
DEFINITION S.pombe chromosome I cosmid c16C9.
ACCESSION 254366
NID 91019812
KEYWORDS MAK32 protein homologue; NAM7 protein homologue; NOT4 homologue;
serine/threonine protein kinase; zinc finger protein.
SOURCE fission yeast.
ORGANISM Schizosaccharomyces pombe
Eukaryotae; mitochondrial eukaryotes; Eumycota; Ascomycotina;
Archaeascomycetes; Schizosaccharomycetales;
Schizosaccharomycetaceae; Schizosaccharomyces.
1 (bases 1 to 14071)
Badcock,K. and Churcher,C.M.
REFERENCE
AUTHORS Barrell,B.G., Rajandream,M.A. and Walsh,S.V.
JOURNAL Unpublished
2 (bases 1 to 14071)
REFERENCE
AUTHORS Barrell,B.G., Rajandream,M.A. and Walsh,S.V.
TITLE Direct Submission
JOURNAL Submitted (12-OCT-1995) to the EMBL/GenBank/DBJ databases.
Schizosaccharomyces pombe chromosome I sequencing project, Sanger
Centre, Hinxton Hall, Hinxton, Cambridge CB10 1RQ E-mail:
barrell@sanger.ac.uk

COMMENT
Notes:
Details of yeast sequencing at the Sanger Centre are available on
the World Wide Web.
(URL, <http://www.sanger.ac.uk/yeastpub/swv/home.html>)
protein coding regions (CDS) have been predicted with the help of
computer analysis using the Genefinder program in Pombase (an ACEDB
database) with additional predictions for the branch-acceptor sites
supplied by the program Sp3splice. CAUTION: It is possible that for
any individual CDS we may have underestimated or overestimated the
number of introns/exons or we may not have chosen the correct
splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC5H10.01c. SP (S.
pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c
(complementary strand).
The more significant matches with motifs in the PROSITE database
are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the
sequenced clone. It may be shorter because we only sequence
overlapping sections once, or longer, because we arrange for a
small overlap between neighbouring submissions.
Cosmid c16C9 is overlapped by cosmid c4G8 at the 5' end and by
cosmid c2G11 at the 3' end.

FEATURES
source
Location/Qualifiers
1..14071
/organism="Schizosaccharomyces pombe"
/strain="972h-"
/chromosome="I"
/clone="cosmid c16C9"
/map="IL"

misc_feature 1..267
/note="overlap with 468"
/complement(96..101)
/note="splice donor sequence, gtaagt"
CDS complement(join(>100..179,217..256))
/gene="SPAC16C9.01c"
/note="SPAC16C9.01c, partial spliced CDS, len: >39,
similar to SW:MK32_YEAST P23060 MAK32 protein (46.7%
identity in 30 aa overlap); pid:e204615"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019813"
/translation="MDYPSKDYEVFTTGMFIIDEIEFDGKKIKDIIGAGSFG"
misc_feature complement(180..193)
/gene="SPAC16C9.01c"
/note="splice branch and acceptor sequence,
ctaagatgcaag"
misc_feature complement(211..216)
/gene="SPAC16C9.01c"
/note="splice donor sequence, gtaaaa"
misc_feature complement(672..682)
/note="splice branch and acceptor sequence, ctaacacttag"
CDS complement(join(777..1037,1137..1697,1747..1848))
/gene="SPAC16C9.02c"
/note="SPAC16C9.02c, unknown, len: 307, similar to
SW:YPE2_RHOKU P23139 hypothetical 25.8 kd protein (46.9%
identity in 128 aa overlap); pid:e204616"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019814"
/translation="MNKQAEPIILGVIGSGFYDLPGFDIVSNPITPGYSPASPIS
IARTSGFLIAFLARHGVTGHIYTPTEVPSRANIALKSLGVLAIVSFAVSLREDIP
PEDEVLPQIIDRTLCARPNTFESEGCVAHVSGDFPDQDLEILSSCGSLKNGSKL
HTKRKGDLDLVVCMGEPASSTRAESNLYRWGASIIINMSYIPEAKIAREAEIAYQMC
MATDYDCWMEPEVTEVMEHISNNKNAKIFLEAVKLEAPLLOGFLGRNLRES
VONGIQTNHKNRPDAIRLQFLFPNLTIPIH"
misc_feature complement(1038..1051)
/gene="SPAC16C9.02c"
/note="splice branch and acceptor sequence,
ttaacaaaaatag"
misc_feature complement(1131..1136)
/gene="SPAC16C9.02c"
/note="splice donor sequence, gtaagt"
misc_feature complement(1698..1713)
/gene="SPAC16C9.02c"
/note="splice branch and acceptor sequence,
ctaagctttattag"
misc_feature complement(1741..1746)
/gene="SPAC16C9.02c"
/note="splice donor sequence, gtagta"
CDS 3172..4668
/gene="SPAC16C9.03"
/note="SPAC16C9.03, unknown, len: 498, similar to
SW:YHX0_YEAST P38861 hypothetical 59.1 kd protein (54.9%
identity in 494 aa overlap); pid:e204617"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019815"
/translation="MDMDSGSMGVAFQPSPIASTILCEGCVPTPPNAAMCMDCIK
MTDITSGIPRESTVNHCRECYMOPNNMMIAPLESRELMALCLKRLGLNOYRLV
DANFIWTEPHSRRIKVLTVQKEFTNTILQOSQOYEVYVNNTOCPDCARTYTPHIWK
AVCQVRQKVLHRTFLYLEQIILKHAHMTVNITKTDGIDFEGORAHAIKMYEEL
SAVPIRYKSEELISEDSEKSNANYKTYSTIEYPICKDDLVCPLTVAKAHGNIAQ
LVVCTKVGPTIRFLDLTLQTCMDLPSIYWRTPPALADIPELTEFIVADVDLGPNTN
GVALADVELIKSSDSTHLTRHLGILNAGNTVLAHLAVTNENNEVDTLAEESI
PEVIVKTYPOTKKNNRNRKLTIGMOKAEDYKQDIERQREYELFLQNLDEDPDEL
RQGVNLKAPYKAIAVADTDEDEDEDIPQISVDLIDVEMHI"
CDS complement(join(5081..6317,6406..6585,6636..6688))
/gene="SPAC16C9.04c"
/note="SPAC16C9.04c, regulatory protein, len: 489, similar
to SW:NOT4_YEAST P34909 general negative regulator of
transcription subunit 4 (NOT4) (38.8% identity in 417 aa
overlap); pid:e204618"

misc_feature /codon_start=1
/product="unknown"
/db_xref="PID:g1019816"
/translation="MLKQEIYSSSEDEDDMCPLCMEIIDSIDKNNKPCQCGYRVCRF
CWHIKEDLNGRCPCACRLYTEENYQWRPVTAEEMKMDLHKNRKRKEKREVEELS
NRKHLANIRVYOKNAYVNGLSPKYANEENINLVKGPPEYFGQYKTIKAIINKAAN
SANGHYGYITYQKEDARAIAIDGVSDDRHLRASYGTTKYCTSYLRNOQCPNS
CMYLHEPGEVDVSTKEDLASLOHTRLSTKPNVNGATHSPSPSLPFTPLPVTKT
PLEANSSPAQNOHITTVDHVHPOVSMTPSLSTNNATSVAPYSSAASVNVPGHA
TTILHEESSALPTAAWAKLSPSVLQERLRAAVNQOPLDALKSSSTQTSIPKIOKIK
AAKLPESEENTTKMINKAINDLVSSLSKINFSTEGTEFDKKQIEMIONLPLFVFNAR
SVLDKEVPEQEKSAENQPTSLGINNGNPVMPPGFQS"
misc_feature complement(6318..6330)
/gene="SPAC16C9.04c"
/note="splice branch and acceptor sequence, ctaatgttttag"
misc_feature complement(6400..6405)
/gene="SPAC16C9.04c"
/note="splice donor sequence, gtagtgt"
misc_feature complement(6586..6605)
/gene="SPAC16C9.04c"
/note="splice branch and acceptor sequence,
ttaactgacaatgcaacag"
misc_feature complement(6630..6635)
/gene="SPAC16C9.04c"
/note="splice donor sequence, gtagtgt"
CDS 7638..8852
/gene="SPAC16C9.05"
/note="SPAC16C9.05, unknown, len: 404, possible zinc
finger protein, similarity to SPAC2F7.07c (in embl entry
250142) 39.1% identity in 115 aa overlap; pid:e204619"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019817"
/translation="MASSINSSQPTVPSISNNSHSDSEFVNEGPPSNFKNNLSSTH
SSFDHVNVLPIQSQKEMDISPVKKQKASINSPKAPLOKSRGSSLSKSHLETESQ
TPVKRRRRKATIRNVDCSAGGRLFCCEGPCSFHLSLEPLTPENIPBEGSMFC
VTSIKSHBPBPKPLSINSLQYDIDSONSQRPLPDDLHYHFGISRGDTGAYKETE
GEMDTDEFSALPTGSSITNAYCGYCSKPSMGACWVGCDLCTFYHKNCKEHAACS
HDSICKKGRVPPKNNAVIRPTLDTSTNLNPKVMISGQFLMGEPSPDELLYFPRL
PVSLCYKVSDEGLIKDFLYAIGIEAKKFNNEKKRELEVIPPDVKSALPARTHPNLP
IALRTLFNKART"
CDS complement(9398..12205)
/gene="SPAC16C9.06c"
/note="SPAC16C9.06c, len: 935, similar to SW:NAM7_YEAST
P30771 NAM7 protein (UPF1 protein) (51.0% identity in 931
aa overlap); contains PS00017 ATP/GTP-binding site motif A
(P-loop); pid:e204620"
/product="unknown"
misc_feature complement(10913..10936)
/gene="SPAC16C9.06c"
/note="PS00017 ATP/GTP-binding site motif A (P-loop)"
CDS 13084..>14071
/gene="SPAC16C9.07"
/note="SPAC16C9.07, len: > 329, partial orf, overlaps
cosmid c2G11 orf SPAC2G11.01, putative serine/threonine
protein kinase; pid:e204621"
/codon_start=1
/product="unknown"
/db_xref="PID:g1019818"
/translation="MVGLISTSEPLPKQESSKSSAPVSNFLSPNNLTQTCSPLRAS
TERKEVFLSQROYSHNSKILTELINSVKRPKNKPPQNEKSAVGEIEKSFDEHLAOK
KGLHFAADLKEIFLNKILSTLDGQDASLHDNIOGKLDNPVISOARRIVLEKLANPN
SKSEYETIPTVENRHPNLNLSKRTLAYEETINDIDETPGGTINPNNKKTLSSETNK
YEPQOSKFHECSQFASPRSSIVGPRTLGKNSKRAADTARMAFRKPSNFNNNIOSS
SYGHAOSSTKLTSGRDNDHQDLNFSYKSIPLNNRPPSPMSLEYFGSGSTTP"
misc_feature 13963..14071
/gene="SPAC16C9.07"
/note="overlap with cosmid c2G11"
BASE COUNT 4537 a 2552 c 2494 g 4488 t
ORIGIN
Query Match 1.5%; Score 35; DB 48; Length 14071;
Best Local Similarity 62.1%; Pred. No. 2.62e-05;

Matches	90;	Conservative	0;	Mismatches	55;	Indels	0;	Gaps	0;
Db 1136	ctgtagccatgcaaccatttgtaggcaatctcagcttcgagctagtgtgctcgg	1195							
Cp 779	ctgtccgcatgCGGATACTGCGTACAAATTCAGCCTCCTTAGCAAGAACCACTCTG	720							
Db 1196	gaatgactgacatgttgataaactgctcccccagcttcgtataaattgcttcgcac	1255							
Cp 719	GAAGTGTGTCATGTTGATTAACATCCGCCCCCAAGTGCAGAACATGAAGCTTTCTGCC	660							
Db 1256	gcgtactaatgacgagacctccat	1280							
Cp 659	GGGAGCTAAAGCAGGTCCCTCGAT	635							
RESULT	4								
LOCUS	CELB0228	38946 bp	DNA		INV		30-MAR-1995		
DEFINITION	Caenorhabditis elegans cosmid B0228.								
ACCESSION	U23168								
NID	9726361								
KEYWORDS	Caenorhabditis elegans strain-Bristol N2.								
SOURCE	Caenorhabditis elegans								
ORGANISM	Eukaryotae; mitochondrial eukaryotes; eukaryote crown group; Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria; Pseudocoelomata; Nematoda; Secernentea; Rhabdilita; Rhabdilitida; Rhabdilitina; Rhabdilitoidea; Rhabdilitidae; Caenorhabditis.								
REFERENCE	1 (bases 1 to 38946)								
AUTHORS	Wilson,R., Alnscoough,R., Anderson,K., Baynes,C., Berks,M., Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J., Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A., Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M., Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N., Latreille,P., Lightning,J., Lloyd,C., McMurray,A., Mortimore,B., O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A., Saunders,D., Showkeen,R., Smaidon,N., Smith,A., Sonhammer,E., Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M., Vaughan,K., Waterston,R., Watson,A., Weinstock,L., Wilkinson-Sproat,J. and Wohlman,P.								
TITLE	2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans								
JOURNAL	Nature	368 (6466),	32-38 (1994)						
MEDLINE	94150718								
REFERENCE	2 (bases 1 to 38946)								
AUTHORS	Leimbach,D.								
TITLE	The sequence of C. elegans cosmid B0228								
JOURNAL	Unpublished (1994)								
REFERENCE	3 (bases 1 to 38946)								
AUTHORS	Waterston,R.								
TITLE	Direct Submission								
JOURNAL	Submitted (21-MAR-1995) Robert Waterston								
COMMENT	Submitted by: Genome Sequencing Center Department of Genetics, Washington University, St. Louis, MO 63110, USA, and Sanger Centre, Hinxton Hall Cambridge CB10 1RQ, England e-mail: tw@nematode.wustl.edu and jesus@sanger.ac.uk								

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is B0495, 8001 bp overlap;3' cosmid is C06A8, 750 bp overlap. Actual start of this cosmid is at base position 4495 of CELB0228; actual end is at 6105 of CELC06A8

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in preparation).

FEATURES	source
CDS	
Location/Qualifiers	
1..38946	
/organism="Caenorhabditis elegans"	
/strain="Bristol N2"	
join(56..502,1135..1583,1735..2017,2219..2342,2390..2842,2988..3076,3122..3241,3287..3355,3404..3472,3519..5054,5181..5318)	
/gene="B0228.4"	
/map="III"	
/note="coded for by C. elegans cDNA yk26d12.5; coded for by C. elegans cDNA yk21b9.5; coded for by C. elegans cDNA yk16f1.5; coded for by C. elegans cDNA yk32h8.5; coded for by C. elegans cDNA yk16f1.3; coded for by C. elegans cDNA yk32h8.3; coded for by C. elegans cDNA yk26d12.3; coded for by C. elegans cDNA yk21b9.3"	
/codon_start=1	
/db_xref="PID:g726370"	
/translation="MNDYETIRDLTGSSQKSNOKISNNGSAHFNQISENVKTESDD DSTCSSLNKLNSNRHSEPMNHLVSGFVAKDLSSLTIPLPQTVIRTCOSIDIDTPDR SRNDKLTSEFYSSCKNLLTTIIPDLVAAGOSTAOLVAQIGMWVCVCCRC CLCKCIANGRGVRGQKQIOEPTIADDDGRDLPYEPGSIHIPKPYNKKLSKSSSS VPHVSVPSLVSSADYKRLSMWETSLNVEDVEPLHTPLVGRKLDYIDKEGNPSC CREKGTTRKPGFRPLRGSGPPPPDPSTSTKVAVTSASEVKKIKDHKKQKKEKKK MDKROSSSGGFFTRWFGSGSNSQNLAEVDIDARSEKTAQOREQLQKESERS GGRTHSHHEEYRRHQPSLWNTFNDEDYATIDYRRSNRMSNREMPASRPNHFVEDT SGPLSRMTNESAFGDRALQYRPAQGAAGPTTGAPTSVNOQDEATLIDLRLSTE PSPVPSRRALPKSASLSVOOKOPITKADGGQLKGSVYTWQNSVDTTDDERAKDF LRGDSSRLSPQSERKNERQIQROQSSGPTNRETEIEYEERKQGPVVRTVEGKL KMERIVGADLITVDSKISSAWTVRDTVTNYKIKSTIGKSLILEIKDQSKYKITLI ENGRTMEREASLDVPDYNKKDYLAEVSKLLSLDREDSVSALTHIEVEIVEDVT NLKTVYIGERADVDVAEQLRLHYEQTADKTPSPIDPLEKVEKIYDEQLDKIELED PEKADHLIKADPLVHTITFLKSKMRQKAGYEMEOGQRFEDETLRLIKRYET DTSNTYKIAKVPRLHVTITFLKSKMRQKAGYEMEOGQRFEDETLRLIKRYET DEEERKQHAVIQHVEEVVATMTKTOREVEAEGQYEMSOEGILRGETFEKRGKHL DSESEERFMEREAEQGOYAMQEGERLLGKFRSKGRHYESESESNASWNGSPT LVDLVKKESSSIFEATFETANNHSPIVAETVRPRPKKENTTIGCTISNQKATSANAE LTKHVNTEKSGSGFRELGEQOAMMLCGFENQSKSEKVEVTSRQOKNETKVAFAAGSAE TIENTTISTTIFADSPAVEGSSNATATYGRKEMESENASNVYLQKSESSSN LSGAETMKHIIQROSSSEARSEFKQVAESCAVMIKNTGVERGSTSTVAEATG"	
join(6960..14594,15067..18774,19220..21034)	
/gene="B0228.3"	
/map="III"	
/codon_start=1	
/evidence=not_experimental	
/db_xref="PID:g726364"	
/translation="MEESWAVECKRRSSSIDRASLKLRAEQEVCYCTGFWNTSGEST RRTILOKTSVESMCLKTKSASSTSIDMFSDCQRAVTSQITSTEIASHREIIVAAGF ISTQSLERLLKFEVIDWENTMNVSOKEQLSANQALTSAPVNLIDPSDLGRIVAPA EQDAYSDLIMRERAESRVFAOLRASADEITIREVALGMSQLEQAAMFILLITSVRC DLRTIAPSNIITANTEIFYVAEEKMSVGMKRESKPKFTISREIIOGFWKG ERDEERVVTKLDRMVSFKSLSLVEAASSTSENVSLGMKADOKSEFTISOKLSPKEI VSEAYGVSESKIDQFFQVMEKMDWSNIELGEKEHVALSANIRTLAPEQVCDGILGKL KAPKQODESVGTQIOEIRATAVMSVRASLLTITSSNSNFSKNISDSEKAVFSNLISV ITSHNSTIGTSETTASISYODIPEMLAASKMLISRNSEKLKKEIREVVIQTVESF WNTTNDQEKVAVLLNEKIDISYSLNTLAASMENELISQELVRHTENACKVNIKPISP RELVSSSEFKITSSDLOQFFNYLEKMDWSQISLPAQEHALISKVRSIAPSASELSNLT GKLRAPEMQEETDLKIRQAQARVVLNVSESMSTISSSETFSRIPENKRAATLTNLV GITTDLSTLGASSENQITEMDYTEAKELELASKRLVDKNLSVLTETRESGDEVQ GFWNTASDKKEVGAIIICEKLSIHHTLHTHAIRTVTEISLTDIOKEOQLSTHFFVKL STRDVYQAAFRISSESVDLTLVEKMDWSKIDLVPEVYSNISSNVQALATSSQSSG ILGKLPPAPEEETTSKEIREINTAKCILNVISSTITSDASLDMISESOAIFSN IIGVASNNTLTSSQSTSSSGFNWVELSEARVLOESNKONLQKYRESSEELIHG IMSTASETEKVAIVKREKLETVHOAMKTLAIOATQSOVNSSELIGSEGNLASIKSITLP TREVISAAFGISNEVIOKMLEVLTKGEWSTISLPOSEYESISQNVRAIAPSPNCDSI LGKINAPGPQTEETTSHELIEKNTAAAVANKAAVESVISKESSLAKLPADAKALITNL AGLVLSKDVSSLCTSSSETFGFDQIRFNQANANISGAVRSETLQRLREFIENQOVGF WSTASNOEKASFIQKQKLDNMYDTMKVIAIOMISQITDGDSSASISSEFTFNIDKAA REVVAEFGIANESVOKALELVNKVEMSDISLPERLQONISSNVRLADSSYNCDSII GKLNAPBQSAIVDQIOSEHOTLEVIANIKSAVSVISNDSVLSKSEDEKAVISNIA NLIISCDLASMSTSNSEFORDILEPONANVILGTPASOVFTRNLOEPIOSEVOGFW STNSQDEKASLVIROKLNVYDKAMKMATETIAYOSLNSNITSEETNLESOKSFKSIR ESIMASFGVSEHVOKTLEITLSKSELSTIVLPAYEATVONIKALSEPNCESIIIG KLNAPDQSAVMDIIMNEQRKLTIVSVNISQSAIESVISNDSSEKLTEDVKSIVILKISE	

```

IVSHDLTALSSSCSDFNLLQESNTQEKADIFLKTIPNSQMLIERLREPIEKOVGFWS
TASNSEKQEMFLKEKVEITIHAMLOTFSASLSEVTVQRDEMATVQSLAARSIOQTPRE
ILCAFGISNQIEQCTFRGLDEVDWSHIDIPCALRDNLVLRIVNADAPNVGNLIS
RPDENQETSTVLEEKNRIQFLNLRSDQTVQMTSSLTRSDERIEVANSIISLIS
ENLDLISOAVOLAQPNMSDETEKSFEPNILLNRVPEETSEESIOSFWKTSQAAE
ASSTISEKLSMLQSEFVSAKQVSTLTDYRRKIFNQNSEIIFGDLTRDVKAAS
VSDETLNTLFSQEQSDMSQIKLSKQKSLSANIKSLATSNLDTLIGQLMPREADSE
OSEASLSERSVSEMENTFKIOQIIFDTVVRKFGSDERASIIINLIGLTSVNLKAAIQ
VASEERLPDLSLSTASITPVFKIQEHYRQPTQOVVQGWKSRHPNETHVEVNRKIEV
LKSILNCYAVAECHEKSLERSLEKPSQODETCATRGVTEVQNAFAVTEQSYSKILKYL
GSMQNFEPFAVKETISKNLKILNPPVDTLVSKYKESHVSKTVEQELVOQMFENK
SALNSEIEMSONSKCIEEETVVLPLNVAIILASANALIVPTESRQSDSENILNISI
QHEESERSAATVRDEINILRHVLOTFAIQKASSEVSRYTEIIKTSQSFHLVHTMLEE
IKTIQLRMEVLKRVSMNOSTAEELLQEAQNSEQOQMETFORIVQNVHWEETKLTETI
RNLRIINMTAIPRESTVGSKAPDQOHSADSYISSTRAEAVINLMAAADNAISTSS
TUSMESDERLHKTILKTMVACNLTPASRSETEMTQGFYOKVOCOCOEQLAERP
KINCEMKCEVETCTGYTEMNCPDAEITTFSSALLGLTSLICDLVTMASSNIQVDSKYD
YRRPAPKNAETTFTESNADAFSALQAGEVTSIGIWSTVSNSDAKTTVSDKLISI
SKTEMSMKAVSENAVNOHSLKSDSIEAVEMNIPDSRKEILOKNYSIDRSNSTVENOG
ATASEDIAIQSAPRVDNVSQTRMSQSQKDLFGSGFGELEPLPQOEDVETTARSR
VYRSSQVRAPSEESIQKTEALRSESLSESNARKTFVDRRENVSLSRKASVERETNM
EARVAKADQSIPIVETLOKSKQESIFSSVESKDLNVCGSMTAKPPIGAKVSIQTK
VEKEVTSATMTVSAVASVECEVGLSEKREOSRDATGSMVRAKSIEVEREFGEVYSTE
KTELEKDOIESWIKSIPQRLVEESAEGFNDEVIIGVMGSLNPLEQOEEETEXLTM
KRSASEARSLKATKESIEEADBEFSKIEADQELITVOKELVKATGILNAAASREINAM
SKMEYSKVPSEVDYIELSLTESRQNSNGOFKETEKEEIVGLMNTGTGENASKYLPHK
PPIDTASMAKAKAKONSIEMTGLQKSPSAEMAGIVSQKTEGANSKFGIAOGAVETT
LTASQSGVTCRDIVSNIGVSETVTOFSXKENGIGFGASNLAPPPESAETFEFTK
ISNLSQTSLNKMAASDVGTYESKIQAPGDNYGDVSLQKVASSDAITKAMQASRDSA
ISVDFRODRETVSAEKSDDLNFKSTNSETOKLKFESKEEESGIFIRSSHEYEEDOKTL
RHRSASRESASRTVAPTNOYQMNFDKKNVEDSVAAEGLSLGIVRESSQSEVQAHER
TSELTKLSMNEEVAGVRAVSEETNETFRGYOQGDVEVQTAGMGRIEAPRPOGEAEI
TOKLRRTLSEYRSKASASEMAESQVTQIQKREDSLASEYVADTLLLKSSYSHVATE
OMSEHLVMSKSEHHISEKQERLIEKESFSSQEFISENOGVHTHWDIDNNGDALIC
WKSSETEOKSLDAKQVTEELTGLDLSVRLGGTDEEKIAEHVIVGASEVAVNISETR
ADMEMSRDADFSDDATYVADLLEERSLSVHEFGESEASTTFGVGLVTKRPEKEV
GRSFSETRKLSQFSDITAISETTDMDSIELRPECDSVTLISHKNOCKDYKDLKAT
VENTADQVTLLEKREVREGADYKKKKEVEDIEEVLPMIRMAEILSKMAKSTVEIQ
AENKLSKPEAOAESQKYLKTANLEINKINVNATHEEVMASATTALECKTSGODVAIRL
RDKSREREKKYQENOMNLSTNAEMETLNLDESVTIAOSVQDSMAFVSKASATVN
LNSDTSIKKAQPEIGIOKSTISQSNVEKTVQFSSATIGNELVTRLGMDLEIEIETLV
EINREQAVGGRIREFGKSEGTGYLYVRALPKVKETSTHTVTATSFROFTSMASG
DEISEANVELTIPSTVGAIEKSTVARSDSMTFSTSHASEYTAADVADADIVASAS
TAARKKAIPTERTSOKLEVGADGIELISLMEGIEITDLDATQOLDLVRKSSIQTIE
SSEEVERINOYLEMPEERGLVTHLINONKETCERNFAYSICSLDTVHSRKSSEIPTY
SVSKVLEKRVLRETWRYIESGDVRFNAVYNLHRYSLSKPTLAQETVLRVVRIAOP
LFISAGDIESDVWSTHOLEKTFVREIEDRTIIGARKGENIKKKFEAGDEKVKLKE
LIGSNLPNEICDISWKIPRLGKVKLETBEFEFDECLFLWSNQLQADR"
join(21201..25567,25621..25699,25774..25874,25950..26055,
26445..26576,26630..26780,26826..26900,26948..27083,
27129..27228,27467..27586)
/gene="B0228.2"
/map="III"
/note="coded for by C. elegans cDNA yk19c1.5; coded for by
C. elegans cDNA yk45a10.5; coded for by C. elegans cDNA
yk21e5.5; coded for by C. elegans cDNA yk21e5.3; coded for
by C. elegans cDNA yk35c1.3; coded for by C. elegans cDNA
yk45a10.3"
/codon_start=1
/db_xref="pid:9726363"
/translation="MRTKESSEDIIVGVIAYNPENLISAIKLDEKASGGCYELTK
AAGDEYKMISSILSKTQAVEKVGKMIQKVAKDEIRVLETTESVANFNYEIPEEN
FKIGIAKTCANNAPEFTMLHECDLFYVMKFVNMKNKDDFESLKITIISNGIESTKL
SCTAAEHIEATRNPENRPAEFSIQKTVVDYINKIEGASLNTFEALNEKMAISTMISR
ODDYSRVEHIVKDKNRGANLKFRIJESSEKQTFPSAFIDTEKEEERTIDLVRRGG
OFKISTDAEENETTLHRDISKPIDTHYDTLHLTLILNSNAPHQIISTSGASNELHTIS
TOLSKPAEWLNTELLIVDKNVEQPTWVRVLECECEVENLHPITYRRPDIIDLDEIWHI
ARNGKFERRCRKAAGDEKEIFEQIKSRGIKDDITDKKFIVNGQEPSTFTTIQISSV
SASVSQDMSRLGEKAIKKVLQNSNKGINVEKKMIATEYRETISEQFRKNDPEFKAD
LTVKDRAGGSYELSTNASEQSSSVSNELICPRPSQLSIEKFTTAQTIIIPALISCK
```

Note: remainder of annotations omitted.

```

Query Match          1.4%; Score 31; DB 25; Length 38946;
Best Local Similarity 65.0%; Pred. No. 6.18e-03;
Matches 67; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Db 33432 catcacttggtttcccatatggtgatctcacgcaactctgactgatacgtlaaataat 33491
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 251 CATCAGATGGCTTGCCAAATGAGTATCCACATATTTTCAGTCTCTCTCTTCTTAAATTT 192

Db 33492 ttgatacttccaaaccactctccccaataatcccaacttgac 33534
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Cp 191 CTGATCATCCAGCGCTGTCTCCACCAGTATTCCAATCTTCAC 149

RESULT 5
LOCUS S73207S1 1117 bp DNA MAM 02-MAR-1995
DEFINITION MHC 67-myosin heavy-chain 67 [exons 31-33, and 40] [cattle,
Genomic, 1117 nt, segment 1 of 2].
ACCESSION S73207
NID 9685132
KEYWORDS 1 of 2
SEGMENT cattle.
SOURCE Bos sp.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1117)
AUTHORS Young,R.B., Hsieh,M.Y., Hudson,J.R.Jr., Richter,H.E. and Scott,M.
TITLE Expression pattern and partial sequence analysis of a fetal bovine
myosin heavy-chain gene
JOURNAL J. Anim. Sci. 72 (4), 903-910 (1994)
MEDLINE 94284221
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI g1b5q 155276] from the original journal article.
This sequence comes from Fig. 3.

FEATURES
source location/qualifiers
1..1117
/organism="Bos sp."
/note="cattle"
join(221..367,455..626,713..837)
/partial

CDS
/gene="MHC 67"
/note="Only first segment of protein sequence is bound;
mismatch(70[K->N]); Description: myosin heavy-chain 67,
MHC 67; Method: conceptual translation supplied by
author. This sequence comes from Fig. 3. Author-given
protein sequence differs from the conceptual translation
for reasons explained in citation. Author-given protein
sequence is in conflict with the conceptual translation. "
/codon_start=1
/product="myosin heavy-chain 67"
/db_xref="pid:g685134"
/translation="KKLAQRLLQDAEEHVAVNAKASLETKORLQNEVEDMLDKKO
RNFDKVLSMKOKYEETQAELELASQKESRSLSTELFKVKNAYEESLDQLETLKRENKN
KNLQOEISDLTEQIAEGKQIHLEKIRKQVEQEKCEIQAALEAE"

BASE COUNT 350 a 243 c 236 g 288 t

Query Match          1.3%; Score 29; DB 34; Length 1117;
Best Local Similarity 70.4%; Pred. No. 8.22e-02;
Matches 50; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Db 894 tggtagaacttgaaccataataacctgataagggcaatttgatgcagtcatacaaa 953
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1350 TGATGGGACTGTAAATGTGTACAGTATTCTTGAGAGGCCAATTTGTAATAATGCATCAAAA 1409

Db 954 gtcttaaaaat 964
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1410 GACTTAAAAAT 1420

RESULT 6
LOCUS HUMRETBLAS 180388 bp DNA PRI 23-NOV-1994
DEFINITION Human retinoblastoma susceptibility gene exons 1-27, complete cds.
ACCESSION L11910
```

NID	g292420
KEYWORDS	nuclear protein; recessive oncogene; retinoblastoma gene; retinoblastoma protein; retinoblastoma susceptibility; tumor suppressor gene.
SOURCE	Homo sapiens DNA.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae.
AUTHORS	1 (sites)
TITLE	Friend,S.H., Bernards,R., Rogelj,S., Weinberg,R.A., Rapaport,J.M., Albert,D.M. and Dryja,T.P.
JOURNAL	A human DNA segment with properties of the gene that predisposes to retinoblastoma and osteosarcoma
MEDLINE	Nature 323 (6089), 643-646 (1986)
REFERENCE	87039336
AUTHORS	2 (sites)
TITLE	Friend,S.H., Horowitz,J.M., Gerber,M.R., Wang,X.F., Bogenmann,E., Li,F.P. and Weinberg,R.A.
JOURNAL	Deletions of a DNA sequence in retinoblastomas and mesenchymal tumors: organization of the sequence and its encoded protein
MEDLINE	[published erratum appears in Proc Natl Acad Sci U S A 1988 Apr;85(7):2234]
REFERENCE	Proc. Natl. Acad. Sci. U.S.A. 84 (24), 9059-9063 (1987)
AUTHORS	88097427
TITLE	3 (sites)
JOURNAL	Lee,W.H., Bookstein,R., Hong,F., Young,L.J., Shew,J.Y. and Lee,E.Y.
MEDLINE	Human retinoblastoma susceptibility gene: cloning, identification, and sequence
REFERENCE	Science 235 (4794), 1394-1399 (1987)
AUTHORS	87149066
TITLE	4 (sites)
JOURNAL	McGee,T.L., Yandell,D.W. and Dryja,T.P.
MEDLINE	Structure and partial genomic sequence of the human retinoblastoma susceptibility gene
REFERENCE	Gene 80 (1), 119-128 (1989)
AUTHORS	90006771
TITLE	5 (sites)
JOURNAL	Sakai,T., Ohtani,N., McGee,T.L., Robbins,P.D. and Dryja,T.P.
MEDLINE	Oncogenic germ-line mutations in Sp1 and ATF sites in the human retinoblastoma gene
REFERENCE	Nature 353 (6339), 83-86 (1991)
AUTHORS	91351319
TITLE	6 (bases 1 to 180388)
JOURNAL	Toguchida,J., McGee,T.L., Paterson,J.C., Eagle,J.R., Tucker,S., Yandell,D.W. and Dryja,T.P.
MEDLINE	Complete genomic sequence of the human retinoblastoma susceptibility gene
REFERENCE	Genomics 17 (3), 535-543 (1993)
AUTHORS	94063891
TITLE	location/Qualifiers
JOURNAL	1..180388
MEDLINE	/organism="Homo sapiens"
REFERENCE	1858..1863
AUTHORS	/bound_molety="Sp1"
TITLE	/evidence=experimental
JOURNAL	1866..1872
MEDLINE	/bound_molety="ATF"
REFERENCE	/evidence=experimental
AUTHORS	<2060..2196
TITLE	/number=1
JOURNAL	join(2060..2196,5424..5550,39446..39561,41926..42045,
MEDLINE	44668..44706,45799..45866,56853..56963,59651..59793,
REFERENCE	61730..61807,64330..64439,65364..65441,70242..70329,
AUTHORS	73753..73869,76430..76486,76889..76920,77001..77077,
TITLE	78083..78279,149998..150116,153208..153353,156693..156838,
JOURNAL	160730..160834,161997..162110,162204..162367,
MEDLINE	170372..170402,173707..173849,174361..174410,
REFERENCE	177005..177078)
AUTHORS	/codon_start=1
TITLE	/product="retinoblastoma susceptibility protein"
JOURNAL	/db_xref="PID:g292421"
MEDLINE	/translation="MPKTPRKTAAATAAAAEPPAPPPPPPEEDPEODSGPEDLPV
REFERENCE	VRLEFEETEEFDFTALCQKLKIPDHVREKRWLTWEKVSSVDGVLGGYIOKKELWIGIC

```

IFIAAVDLDEMSFTFTELOKNIIEISVHKFEFNLKLEIDISTKVDNAMSRLKKYDVLFA
LEFSKERTCELLIYLTQPPSSSISTEINSALVTKVSWTFLLAKGEVLQMEDDLVISFQL
MCLVDYFIKILSPPMILKEPKTAVIPINGSPRTPRGQRSARIAKQLENDTRIIEV
LCKEHCNIDEVKNVFKNFIPFMSLGLVTSNGLEVENLSKRYEEIYLKKNDDAR
LFLHDKTLQTDSDISDFEOTPRKSNLDEEVNVIPEHTPVRTVMTIQQLMILNSA
SDQSENLISYFNNECTVNPRESILKRVKDIGYIFKEKFAKAVGQCGVEIGSORYKLGV
RLYRVMESMLKSEERLSIONFSKLLNDNIFHMSLLACALEVMATYSRSTSONLDS
GTDLSFPWILNLNLKAFDEYKVIIESFIKAEGNLTRPMIKHLERCEHRIMESLAWLSD
SPLEDLIKQSKDREGPTDHEASACPUNLPLQNNHTAADMYLSVRSPPKKGSTTRVNS
TANAEQATSAFOTOKPLKSTSLSEYKKVYRLAYRLNLTICERLSEHPELEHIWTV
LEQHTIQNEYELMRDRLDQIMCMSMYGICKVKKNIDLKFIIVTAYKDLPHAVQETFRK
RVLKKEEYDSIIVEYNSVFQRLKTNILQYASTRPTLSPIPHIPRSPPKFPSSPLR
IPGGNYIISPLKSPYKISEGPTPTKMTPRSRLVSGESFGTSEKFOKINQWNCNSDD
RYLKRSALGSGNPPKPLKKLRPDIEGSDADGSKHLPGESKFQOKLAEMTSTRIRMQOK
KMNDSDTSNKEEK"
2197..5423
/number=1
intron
2302
/standard_name="Restriction Fragment length Polymorphism"
/notes="BamHI RFLP"
5424..5550
/number=2
5551..39445
/number=2
14868..14914
/standard_name="microsatellite RB12 (AFM058xd6)"
39446..39561
/number=3
39562..41925
/number=3
41926..42045
/number=4
42046..44667
/number=4
43218..43257
/standard_name="microsatellite RB14"
44668..44706
/number=5
44707..45798
/number=5
45799..45866
/number=6
45867..56852
/number=6
56853..56963
/number=7
56964..59650
/number=7
59651..59793
/number=8
59794..61729
/number=8
61730..61807
/number=9
61808..64329
/number=9
64330..64439
/number=10
64440..65363
/number=10
65364..65441
/number=11
65442..70241
/number=11
70242..70329
/number=12
70330..73752
/number=12
73753..73869
/number=13
73870..76429
/number=13
76430..76486
exon

```

intron	/number=14 76487..76888
exon	/number=14 76889..76920
intron	/number=15 76921..77000
exon	/number=15 77001..77077
intron	/number=16 77078..78082
exon	/number=16 78083..78279
intron	/number=17 78280..149997
variation	/number=17 99426
satellite	/standard_name="Restriction Fragment Length Polymorphism" /note="XbaI RFLP" 102921..102959
repeat_region	/standard_name="microsatellite RB117" 123912..125501
exon	/standard_name="Variable Number Tandem Repeats (VNTR)" /rpt_type=tandem 149998..150116
intron	/number=18 150117..153207
exon	/number=18 153208..153353
intron	/number=19 153354..156692
exon	/number=19 156693..156838
intron	/number=20 156839..160729
satellite	/number=20 156895..156976
exon	/standard_name="microsatellite RB1.20" 160730..160834
intron	/number=21 160835..161996
exon	/number=21 161997..162110
intron	/number=22 162111..162203
exon	/number=22 162204..162367
intron	/number=23 162368..170371
exon	/number=23 170372..170402
	/number=24

Note: remainder of annotations omitted.

Query Match	1.38;	Score 29;	DB 60;	Length 180388;
-------------	-------	-----------	--------	----------------

Matches	83;	Conservative	0;	Mismatches	48;	Indels	1;	Gaps	1;
---------	-----	--------------	----	------------	-----	--------	----	------	----

Db 7263 ctgttcaataatgaatgaatctctgattatttattagaataattctataag 7322

CP 1509 CTTTATATATAATTCTTTGTATGCATCTCTGATATGGTTTTAAAGAGAAATTGCTAGATG 1450

Db 7323 ttgaattactaa-atcaaaagctatgaacatttttatgactcttgaatgaatattgtctaga 7381

CP 1449 TAGAGTTCACGACCAAGAGAGCTCGATTTTTAAGTCTTTTGATGCATTTTACCAA 1390

Db 7382 ttacacttcaga 7393

CP 1389 TTGCCCTCCAGA 1378

RESULT 7

LOCUS HS117E6F

```

DEFINITION H.sapiens Cpg DNA, clone 117e6, forward read cp9117e6.ftla .
ACCESSION Z58950
NID 91030863
KEYWORDS Cpg island.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
REFERENCE 1 (bases 1 to 127)
AUTHORS Macdonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-1995) to the EMBL/GenBank/DBJ databases. The
Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail
contact: humpub@sanger.ac.uk
2 (bases 1 to 127)
AUTHORS Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
TITLE Purification of Cpg islands using a methylated DNA binding column
JOURNAL Nature Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT Vector: pGEM-5zf(-).
FEATURES Location/Qualifiers
source 1..127

```

BASE COUNT	ORIGIN
52 a	28 c
14 g	31 t
2 others	

Query Match	Score	DB	Length
1.28;	27;	50;	127;

Matches	39;	Conservative	0;	Mismatches	14;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

Db 73 antgaaaattggtacaactttctgagacantttagcaaaaaacacccaaa 125

QY 1357 ACTGTAATTTGGTACAGTATTTCTGGAGGCCAATTTGGTAAATGCATCAAAA 1409

RESULT

LOCUS	HS117E6R	127 bp	DNA	PRI	19-OCT-1995
DEFINITION	H. sapiens cnc	DNA	31000-11706	2000000000	19-OCT-1995

ACCESSION	258951
NID	91030864

KEYWORDS CPG 1s
SOURCE human.

ORGANISM Homo sapiens
Eukaryotae;

Homo.
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Hominidae;

1 (pages 1 to 127)
MacDonald, M., Huckle, E., Wilkinson, P. and Micklem, G

JOURNAL

contact: humpub@sanger.ac.uk

AUTHORS Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
TITRE Duration of the Cretaceous

JOURNAL
MEDLINE
Nature Genet. 6 (3), 236-244 (1994)
94282070

```

COMMENT      Vector: PGEM-5Zf(-).
FEATURES             .   Location/Qualifiers

```

Source

```
/dev_stage="adult"
/sex="male"
```

```

/tissue_type="blood"
/clone_id="CGI-1"

```

BASE COUNT		/clone="11/ab"	
30 a	14 c	29 d	53 t
		1 others	


```

/strain="Landsberg erecta"
/dev_stage="vegetative"
/tissue_type="leaf"
/clone_lib="CDNA in lambda ZAP11"
/clone="PM25"
30..1592
CDS

```

```

/EC_number="3.4.11.1"
/EC_number="3.4.11.2"
/codon_start=1
/product="leucine aminopeptidase"
/db_xref="PID:g16394"
/translation="MAHTLGLTPNSTEPHKISFTAKEIDVIEWKDILLVGVTEKDL
ADGNSPVEPILSKEDAHLSGLLAQVSEEDFTGPGOSVLRPLGLSKRTALIGL
GQSVSSPVAFHSLGEAVATVSKASQSTSAIVLASSVDEKSLSSVALASGYLGLF
EGGRYKSESKPSLKAVDIIGFTGAEYERKLLKVAEDVSYGVIFGRELINSPANVLP
AVLAEAAKAVASTVSUFTVATNINLEEOCKELKMSGYLAAVASANPPIHTHYKKPN
GSYKTKLALVKGKLTDPDSCGYNIKITGBCSIELMKFDMGSSAVALGAARATSEIKPPG
VEVHFIVAACENMISGTGMRPGDVIITASNGKTIENVNTDGEKRLTIADALVACNOGV
DKRVDLATLGCVALGTSMAGITPSELAKEVIAASESGEKLWRMPLDESYWEM
KSGVADVMNTGGRAGCSITALEFLKDFVSEKVQMMHIDMAGPVWNEKKKSGTGFVA
LLVEWVQKNSSS"

```

polyA_signal	1661..1666		
polyA_signal	1734..1739		
polyA_site	1754		
BASE COUNT	483 a	337 c	455 g
ORIGIN			498 t

Query Match	1.28;	Score 27;	DB 42;	Length 1773;
Best Local Similarity	82.98;	Pred. No. 9.78e-01;		
Matches	34;	Conservative	0;	Mismatches 7;
			Indels 0;	Gaps 0;

```

Db      1731 ttgaataaagtgtgctactgttttaaaaaaaaaaaaaaa 1771
        || ||||| || || || ||||| |||||
Qy      2220 TTATAAAGTGAAGCTGCTTTTAAAAAANAANAANA 2260

```

RESULT	12			
LOCUS	S67474S5	2126 bp	DNA	PRI 17-FEB-1994
DEFINITION	(GC+IS)-vitamin D-binding protein/group specific component [human, Peripheral blood leukocytes, Genomic, 2126 nt, segment 5 of 9].			
ACCESSION	S67479			
NID	9455965			
KEYWORDS				
SOURCE	5 of 9			
SEGMENT	human peripheral blood leukocytes.			

ORGANISM	REFERENCE
homo sapiens	Unclassified.
1	(bases 1 to 2126)

REFERENCE	1 (bases 1 to 2126)
AUTHORS	Braun,A., Kofler,A., Morawietz,S. and Cleve,H.
TITLE	Sequence and organization of the human vitamin D-binding protein gene
JOURNAL	Biochim. Biophys. Acta 1216 (3), 385-394 (1993)
MEDLINE	94092730
REMARK	GenBank staff at the National Library of Medicine created this

FEATURES	
source	Location/Qualifiers
	1..2126

BASE COUNT	695 a	356 c	345 g	730 t
ORIGIN				

Query Match	1.28;	Score 28;	DB 62;	Length 2126;
Best Local Similarity	78.0%;	Pred. No. 2.88e-01;		
Matches	46;	Conservative	0;	Mismatches 12; Indels 1; Gaps 1;

D_b 1156 attttaagacatttaagtacataatgacaactgcctttcc-aaatatgtaccaattt 1213
||||| |||| | ||| |||| | |||| |||||
C_p 1420 ATTTTAAAGCTTTTGATGCATTTTACCAAATTGCCCTCCAGAATACTGTACCAATTT 1362

RESULT 13

LOCUS	DDISASIA	2145 bp	mRNA	PLN	15-SEP-1990
DEFINITION	D.discoidium GTP-binding protein (sas1) gene, complete cds.				
ACCESSION	M34456				
NID	g167874				
KEYWORDS	GTP-binding protein.				
SOURCE	D. discoidium, cDNA to mRNA.				
ORGANISM	Dictyostellium discoidium				
	Eukaryota; Fungi; Myxomycota; Acrasiomycetes; Dictyosteliales;				
	Dictyostelaceae.				
REFERENCE	1 (bases 1 to 2145)				
AUTHORS	Saxe,S.A. and Kimmel,A.R.				
TITLE	Sas1 and Sas2, GTP-binding protein genes in Dictyostellium discoidium with sequence similarities to essential genes in Saccharomyces cerevisiae				
JOURNAL	Mol. Cell. Biol. 10, 2367-2378 (1990)				
MEDLINE	90220623				
FEATURES	Location/Qualifiers				

```
source      1..2145
/organism="Dictyostelium discoideum"
CDS         1095..1706
```

BASE COUNT	929 a	249 c	238 g	727 t	2 others
ORIGIN	/translation="MTSPATNKPAAAYDLKLLLIGDSGAGKSCILLRFSDSEFTPSF IATIGIDEFKIRITIEEGKRILKQIWDTPAGQERFRTITAYRGAMGILLYDVYDEKS FGSIRNWRIRNEOHASDSVNKLIGNKDMTEKKNVSDSGKSLADEYGIKPLETSAK NSVAYEEAFIGLAKDIKKRMIDTPNDPDHTICITPNNKNTCC"				

Query Match	1.28;	Score 27;	DB 43;	Length 2145;
Best Local Similarity	81.48;	Pred. No. 9.78e-01;		
Matches	35; Conservative	0; Mismatches	8; Indels	0; Gaps 0;

Dd 1937 taataataataataataataataacaaatttaa 1979
||| ||| ||| ||| ||| ||| ||| |||
Qy 1519 TAATAATGATAGTTATATATAAANAATTGAACAATCTGAA 1561

[illegible]

REFERENCE
1 (bases 1 to 3030)
Eukaryota; Animalia; Protozoa; Apicomplexa; Sporozoa; Coccidiasina; Eucoccididia; Haemosporina; Plasmodiidae.

TITLE	JOURNAL
The expression pattern of <i>Plasmodium falciparum</i> suggests a common mode of gene regulation during bloodstage development	unpublished (1993)

FEATURES	Location/Qualifiers
source	1..3030

```

/organism="Plasmodium falciparum"
/dev_stage="bloodstage"
/sequenced_mol="cDNA to mRNA"
1122 a      269 c      397 g      1242 t
BASE COUNT
ORIGIN
```

Query Match	1.28;	Score 27;	DB 30;	Length 3030;
Best Local Similarity	72.18;	Pred. No. 9.78e-01;		
Matches	44;	Conservative	0;	Mismatches 17;
			Indels	0;
			Gaps	0;

D_b 280 aaatgtctataaaaaaaataaataataatataatataatattat 339
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 QY 1727 AATGCTATGTAATATATCAAAAAAACTAGAAACAATATATATACTGTATTCAT 1786

Db 340 t 340

QY 1787 T 1787

RESULT	15
LOCUS	RRCYTBC1 5550 bp DNA BCT 15-FEB-1991
DEFINITION	R. rubrum cytochrome bcl-complex genes (petA, petB, petC).
ACCESSION	X55387
NID	946382
KEYWORDS	cytochrome b; cytochrome bc-l complex; cytochrome c1; petA gene; petB gene; petC gene; Rieske Fe-S protein.
SOURCE	Rhodospirillum rubrum
ORGANISM	Eubacteria; Proteobacteria; alpha subdivision; Rhodospirillum.
REFERENCE	1 (bases 1 to 5500) Majewski,C. Direct Submission Submitted (05-NOV-1990) to the EMBL/genbank/DDBJ databases. C. Majewski, ROHR UNIVERSITAET BOCHUM, LEHRSTUHL F BIOCHEMIE DER PFLANZEN, ROHR UNIVERSITAET, POSTFACH 102148, 4630 BOCHUM 1 2 (bases 1 to 5550) Majewski,C. and Trebst,A. The pet genes of Rhodospirillum rubrum: cloning and sequencing of the genes for the cytochrome bcl-complex Mol. Gen. Genet. 224 (3), 373-382 (1990)
JOURNAL	91094774
AUTHORS	
TITLE	
JOURNAL	
MEDLINE	
FEATURES	
source	Location/Qualifiers 1..5550 /organism="Rhodospirillum rubrum" /strain="FRI" /clone_lib="Partially digested Sau3A fragments of genomic Rs. rubrum DNA in EMBL3." /clone="EMBL-pet 11.1, EMBL-pet 4.1" <3..>770 /gene="ORF1" 3..770 /gene="ORF1" /codon_start=1 /transl_table=11 /db_xref="PID:g46383" /translation="MLDAPAGRWDPDAGEGALRLLDGEKDLVIDRALLRMAERG KGYPRAERTARWRALLETALGTATPTPLGGDALACGLKPAYGEAMRPCADAGLR AAAPCGTRCWPGSPIGGTGPTRNRPBSRAMNRYSRSDRPSGVLRPACGMAGLAEL HOGGDGDERAHAGNHNRNPDAQVGALVDLGNNARLSAQOQHVVGLKVSFGMKRARL RCQQDAQGFVGEPRDEIRPGRMTHKSFAFEIVHSGA"
mRNA	<3..>770 /gene="ORF1" 1250..1253 /product="Rieske Fes protein"
CDS	/db_xref="PID:g46384" /translation="MAEAHTASTPGESSRRDFLIYGTAVGAVGVALAVWPFIDFM NPADTLALASTEVDVSAIAEGQAITYWRGKPVFVRHRTOKEIIVARA VDPASLRDP QTDEARVQAQAWLVMGVCETHLGCIPLGOKAKGDPKGDFDGWFCPCGHSHYDSAGRIRK GPAPLNVPVPYAFTDDTVLLIG" 1811..1814 /db_xref="PID:g46385" /translation="MTTPPRWNNAKLMFDERLPVLTVAHKELVVPAPRNLYFWNF GSLAGIAMITIMATGI FLAMS YTAHDHAFDSVERIMRDVNYGWLRYMHANGASMFF IVVYVMFERGLYYGSYKPPREVLWLTGLVILLMATATAFMYGLPMGOMSF GATVIT NLFSAIPIVVGDDIVTLLMGESVDNPTLNRFSLHYLF PMLLEAVVF LHMWALHVKS
prim_transcript	
RBS	
prim_transcript	
CDS	

DB	4387	atcgccgatacgttcacgcaccgcgacc	4413
DB	4327	atcttgcgtctccgcgcgtcgttctctc	4386
QY	386	gtcatagtgaaccacagcttctgagggag	445
DB	4267	ccctccgatagtgaactaccgcgcacata	4326
QY	326	ccttcaaggctacactaccgacacacat	385
DB	4327	atcttgcgtctccgcgcgtcgttctctc	4386
QY	386	gtcatagtgaaccacagcttctgagggag	445
DB	4387	atcgccgatacgttcacgcaccgcgacc	4413

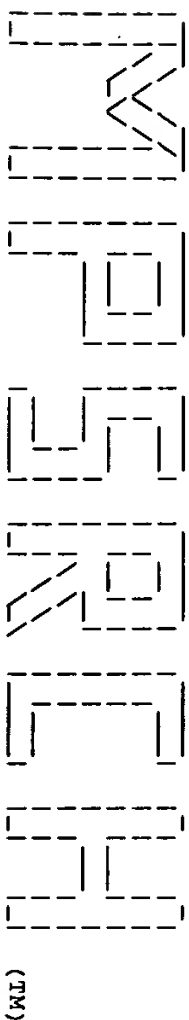
Tue Feb 11 11:27:24 1997

US-08-674-311-1-01.rge

Page 12

OY 446 ATTATGATCAGTTCATTCATGACAGACC 472

Search completed: Mon Feb 10 22:51:01 1997
Job time : 1616 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

MPsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 22:51:20 1997; MasPar time 223.13 Seconds
886.969 Million cell updates/sec
Tabular output not generated.

Title: >US-08-674-311-1
Description: (1-2269) from US08674311.seq (1 of 2)
Perfect Score: 2269
N.A. Sequence: 1 GAATTCGGCTCCGACACTGCT.....AAAAAAGGCGAATTC 2269
Comp: CTTAAGGCGAGCGCTGACGA.....TTTTTTTTTTCGCCCTAAG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22

Statistics: Mean 9.967; Variance 6.722; scale 1.483

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Pred. No.
1	467	20.6	2784	16	Q99202 Pseudomonas putida me	1.22e-267
2	466	20.5	2763	19	T15167 Methylthioadenosine-p	5.20e-267
3	404	17.8	2763	15	Q92813 Human MTase.	6.16e-228
4	94	4.1	1047	2	Q10572 Human Natriuretic Pep	5.36e-37
5	76	3.3	1047	2	Q10572 Human Natriuretic Pep	1.15e-26
6	42	1.9	91	9	Q51746 Oligonucleotide probe	2.85e-08
7	44	1.9	91	9	Q51746 Oligonucleotide probe	2.76e-09
8	41	1.8	204	1	N81164 Base substituted E.co	9.08e-08
9	41	1.8	204	1	N81164 Base substituted E.co	9.08e-08
10	34	1.5	114	12	Q70470 Generic DNA sequence	2.35e-04
11	35	1.5	3018	21	T31452 Human Ndr serine/thre	7.88e-05
12	32	1.4	114	12	Q70467 Generic DNA sequence	2.02e-03
13	31	1.4	114	12	Q70469 Generic DNA sequence	5.82e-03
14	31	1.4	114	12	Q70472 Generic DNA sequence	5.82e-03
15	31	1.4	114	12	Q70465 Generic DNA sequence	5.82e-03
16	31	1.4	114	12	Q70469 Generic DNA sequence	5.82e-03
17	30	1.3	114	12	Q70465 Generic DNA sequence	1.65e-02

18	30	1.3	114	12	Q70468	Generic DNA sequence	1.65e-02
19	30	1.3	114	12	Q70473	Generic DNA sequence	1.65e-02
20	29	1.3	114	12	Q70466	Generic DNA sequence	4.63e-02
21	30	1.3	114	12	Q70467	Generic DNA sequence	1.65e-02
22	27	1.2	66	21	T13585	TSAR-9 library genera	3.47e-01
23	27	1.2	69	21	T13583	TSAR-9 library genera	3.47e-01
24	27	1.2	114	12	Q70471	Generic DNA sequence	3.47e-01
25	28	1.2	114	12	Q70466	Generic DNA sequence	1.28e-01
26	28	1.2	114	12	Q70468	Generic DNA sequence	1.28e-01
27	28	1.2	114	12	Q70470	Generic DNA sequence	1.28e-01
28	25	1.1	74	21	T13613	DC43 TSAR library gen	2.42e+00
29	25	1.1	75	21	T13612	DC43 TSAR library gen	2.42e+00
30	25	1.1	81	21	T13611	DC43 TSAR library gen	2.42e+00
31	25	1.1	82	21	T13610	DC43 TSAR library gen	2.42e+00
32	25	1.1	111	17	T02821	MS-associated retrovi	2.42e+00
33	25	1.1	114	12	Q70472	Generic DNA sequence	2.42e+00
34	24	1.1	114	12	Q70473	Generic DNA sequence	6.22e+00
35	25	1.1	501	3	N50033	Sequence encoding new	6.22e+00
36	24	1.1	679	12	Q74410	Lipid transfer protei	6.22e+00
37	26	1.1	1044	2	Q12840	Variable region of Ia	9.25e-01
38	24	1.1	1259	3	Q23028	Clone W264 encoding c	6.22e+00
39	26	1.1	1419	3	Q14664	Tomato cellulase gene	9.25e-01
40	25	1.1	1626	14	Q81066	Insect diuretic hormo	2.42e+00
41	24	1.1	1633	21	T09258	Human ara-4 protein c	6.22e+00
42	24	1.1	2214	11	Q65604	Porcine zona pellucid	6.22e+00
43	26	1.1	4565	1	Q03704	Gene encoding plant e	9.25e-01
44	26	1.1	4565	6	Q36024	Leaf-1 genomic clone	9.25e-01
45	24	1.1	19307	20	T27558	Shuttle vector paddel	6.22e+00

ALIGNMENTS

RESULT	1	
ID	Q99202	standard; DNA; 2784 BP.
AC	Q99202;	
DT	07-MAR-1996	(first entry)
DE	Pseudomonas putida methylthioadenosine-phosphorylase DNA.	
KW	Enzyme; MTase; L-Met-L-deamino-gamma-mercaptopmethane-lyase;	
KW	chemotherapy; cancer therapy; methionine starvation; ss.	
OS	Pseudomonas putida.	
FT	key	Location/Qualifiers
FT	CDS	1..2763
PN	/*tag= a	
PD	WO9517908-A1	
PF	06-JUL-1995.	
PR	22-DEC-1994; U14919.	
PR	29-DEC-1993; US-176413.	
PA	(REGC) UNIV CALIFORNIA.	
PI	Carson DA, Nobori T;	
DR	WPI; 95-246192/32.	
PT	Selective methionine starvation of methyl:thio:adenosine	
PT	phosphorylase negative tumour cells - used in chemotherapy of	
PT	mammalian malignant cells.	
PS	Disclosure; Page 27-28; 46pp; English.	
CC	Antigenic peptides may be generated from this sequence encoding	
CC	methyladenosine-phosphorylase, which are then used to generate	
CC	antibodies specific for MTase. The produced antibodies may be	
CC	used in an immunoassay for the detection of MTase.	
SQ	Sequence 2784 BP; 781 A; 505 C; 598 G; 804 T;	
	Query Match	20.6%; Score 467; DB 16; Length 2784;
	Best Local Similarity	97.8%; Pred. No. 1.22e-267;
	Matches 492; Conservative	1; Mismatches 6; Indels 4; Gaps
Db	2270	agaatattgcccagttctgtttattaccagaacattaaagtacatgctgcccagg 2329
QY	933	AGAAATATGCGCCAGTTTCTGTTTATACCAAGACATTAAGTAGCATGGCTGCCAGG 992
Db	2330	agaaaagaagacattcttaattccagtcattttgggaattcctgcttaacttgaanaaat 2389
QY	993	AGAAAGAGAGACATTCATTCACAGTCATTTGGGAATTCCTGCTTAACCTGAAAAAAT 1052
Db	2390	atgggaaagacatgcagcttccatgccttgcctatcaagaagatgtgttaagaagaagac 2449

✓ this ref only data fine last base 87

✓ this reference
only deleted
line last 39
base
of MTase

|||||
QY 1053 ATGGGAAGACATGCTTTCATGCTTGGCTATCAAGAGTATGTTGAAGAAC 1112
Db 2450 aagacattgtgtat-agagactcctcaatgatttagaacaactcacaatacagaagaa 2508
QY 1113 AAGACATTGTGTATTAAGACACTCTGATGATTAGACAACCTCAAAATACAGAGAA 1172
Db 2509 aagcaaatgactagtaa-catgtggaaaaaataatcatttgaaggggaaaaaaa-- 2565
QY 1173 AAGCAATGACTAGTAACATGTGGGAAAAATATATACATTTAAGGGGAAAAAAA 1232
Db 2566 ccccaaccattctctctcccccatttaatttgcaacaataaagggtggaggtatctc 2625
QY 1233 CCCACCATTTCTCTCTCCCTATTAATTTGCAACAATAAGGGTGAAGGTAATCTC 1292
Db 2626 tacttccatactgcgaagaatgtggaagaagaatggactcttggtattattga 2685
QY 1293 TACTTCTCTATCTGCAAGAATGTGAGGAAGAAATGGGACTCTTGCTTATTATGA 1352
Db 2686 tgcgaactgtaaatgtgtacagatattcttggagggaatttgtaaaatgcatacaaaagac 2745
QY 1353 TGGCACTGTAATTTGTTACAGTATTCTTGAGGGCAATTGGTAAATGCATCAAAAGAC 1412
Db 2746 ttaaaatacggagcgtacnnmr 2768
QY 1413 TTAATAATACGAGCTCCTTTGG 1435

only
in
base
MITP

RESULT 2
ID T15167 standard; DNA: 2763 BP.
AC T15167;
DE 29-JUN-1996 (first entry)
DE Methylthadenosine-phosphorylase gene.
KW Human; methylthadenosine-phosphorylase; chromosome walking;
KW Interferon-alpha; CDK4I; tumour suppressor; chromosome-9p21;
KW cyclin-dependent kinase-4-inhibitor; probe; primer; mutation;
KW leukaemia; melanoma; diagnostic; dysplastic nevus syndrome; glioma;
KW non-small cell lung carcinoma; cancer; gene therapy; antisense;
KW ribozyme; antibody; imaging; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 254..421
FT /*tag- b 422..615
FT Intron
FT /*tag- c 616..720
FT exon
FT /*tag- d 721..963
FT Intron
FT /*tag- e 964..1203
FT exon
FT /*tag- f
PN WO9528169-A1.
PD 26-OCT-1995
PF 12-APR-1995; U04655.
PR 14-APR-1994; US-227800.
PA (REGC) UNIV CALIFORNIA.
PI Carson DA, Nobori T;
DR WPI: 95-373630/48.
PT Cyclin dependent kinase inhibitor gene, related vectors and
PT antibodies - useful for diagnosis, assessing predisposition and
PT treatment of cancers
PS Example 1; Page 96-101; 129pp; English.
CC The sequence encodes a methylthadenosine-phosphorylase, and is
CC located at chromosome-9p21. A cyclin-dependent
CC protein-kinase-4-inhibitor (CDK4I) tumour suppressor gene
CC (T15157-58) is located between this gene and an interferon-alpha
CC gene cluster, and has been isolated by chromosome walking. The
CC CDK4I gene, probe and primer derivatives and the gene product may
CC be used in diagnosis of cancer, particularly melanoma (especially
CC dysplastic nevus syndrome), glioma, non-small cell lung carcinoma
CC or leukaemia. The gene may also be used in cancer gene therapy, or
CC in antitumour antisense oligonucleotide or ribozyme construction.
CC Antibodies against CDK4I may be used in diagnosis or in vivo imaging.

SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T;
Query Match 20.5%; Score 466; DB 19; Length 2763;
Best Local Similarity 99.0%; Pred. No. 5.20e-267;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 agaataggcccaagtttctgttattatccaagaacataaagtagcatggtcccaag 2329
QY 933 AGAATATGGCCAGTTTCTGTTTATTATCAAGACATTAAAGTAGCATGGCTGCCAGG 992
Db 2330 agaaagaagacattcctaattccagtcatttgggaattcctgcttaacttgaaaaaat 2389
QY 993 AGAAAAGAAGACATTCTAATTCAGTCATTTTGGGAATTCCTGCTTAACCTGAATAAAT 1052
Db 2390 atgggaagaacatgcagcttccatgccccttgccctatcaagaagatgtgtgaagaagac 2449
QY 1053 ATGGGAAGACATGACGCTTTCATGCCCTTGCCCTATCAAGAGTATGTTGAAGAAAGAC 1112
Db 2450 aagacattgtgtat-agagactcctcaatgatttagaacaactcacaatacagaagaa 2508
QY 1113 AAGACATTGTGTATTAAGAGACTCTGATGATTAGACAACCTCAAAATACAGAGAA 1172
Db 2509 aagcaaatgactagtaa-catgtggaaaaaataatcatttgaaggggaaaaaaa-- 2565
QY 1173 AAGCAATGACTAGTAACATGTGGGAAAAATATATACATTTAAGGGGAAAAAAA 1232
Db 2566 ccccaaccattctctctcccccatttaatttgcaacaataaagggtggaggtatctc 2625
QY 1233 CCCACCATTTCTCTCTCCCTATTAATTTGCAACAATAAGGGTGAAGGTAATCTC 1292
Db 2626 tacttccatactgcgaagaatgtggaagaagaatggactcttggtattattga 2685
QY 1293 TACTTCTCTATCTGCAAGAATGTGAGGAAGAAATGGGACTCTTGCTTATTATGA 1352
Db 2686 tgcgaactgtaaatgtgtacagatattcttggagggaatttgtaaaatgcatacaaaagac 2745
QY 1353 TGGCACTGTAATTTGTTACAGTATTCTTGAGGGCAATTGGTAAATGCATCAAAAGAC 1412
Db 2746 ttaaaatacggagcgt 2761
QY 1413 TTAATAATACGAGCT 1428
RESULT 3
ID Q92813 standard; DNA: 2763 BP.
AC Q92813;
DE 05-DEC-1995 (first entry)
DE Human MTase.
KW MTase; methyladenosine-phosphatase; malignancy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 254..421
FT /*tag- a 616..720
FT exon
FT /*tag- b 964..1203
FT exon
FT /*tag- c 1640..1762
FT exon
FT /*tag- d 2272..2310
FT exon
FT /*tag- e
PN WO9518233-A.
PD 06-JUL-1995.
PF 22-DEC-1994; U14920.
PR 29-DEC-1993; US-176855.
PA (REGC) CIBA GEIGY CORP.
PI (REGC) UNIV CALIFORNIA.
PI Carson DA, Nobori T, Takabayashi K;
DR WPI: 95-246398/32.
PT Detecting methyladenosine phosphatase in mammalian cells - by
PT hybridisation with specific oligonucleotide for detecting malignancy,
PT also new nucleic acid, expression vectors, derived polypeptide(s) and
PT antibodies

CP 216 TTTTCAGTCTCTCTCTCTAAATTTCTGATCATCCAGCCCTGTCTCCACCAATTATTCGA 157
DB 888 dnytcndannndndvykvtntngdaymvvsngnrgnrgnrrhannarmanandavssnrr 947
CP 156 ATCTTCACGGCGGTGTGTGTGTGTGCACAGGCCATGTCTGCACGGGAAAGAATCTGCAG 97
DB 948 hrhndnrrngvhtgnvcagvvgknmyrcngdvtntasrmnsngnanknhvsstk 1007
CP 96 TGGCGAGCGCTCGGACTAAGGGAACCAAGCCAGACAGACGCGGTGTGTGCACACC 37
DB 1008 andngcnmnnrgdv 1023
CP 36 TCACTGCGCGGAGTG 21

RESULT 6
ID 051746 standard; cDNA; 91 BP.
AC 051746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis; ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (051735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also 051735-45 and 051747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;
Query Match 1.9%; Score 42; DB 9; Length 91;
Best Local Similarity 7.4%; Pred. No. 2.85e-08;
Matches 4; Conservative 44; Mismatches 6; Indels 0; Gaps 0;
DB 11 ssvhsyvvvnhshsvhvhvhsvvvvhvvhvhyhvsctca 64
OY 825 GGGTCTTAAGACCTCGAAGAAACGCTAATAAGCCAAAGCTTACTGCTCA 878

RESULT 7
ID 051746 standard; cDNA; 91 BP.
AC 051746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis; ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-378844/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (051735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may

CC be useful as an initial screen for mycobacterial infection.
CC See also 051735-45 and 051747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;
Query Match 1.9%; Score 44; DB 9; Length 91;
Best Local Similarity 2.0%; Pred. No. 2.76e-09;
Matches 1; Conservative 46; Mismatches 3; Indels 0; Gaps 0;
DB 12 svhsyvvvnhshsvhvhvhsvvvvhvvhvhyhvsvc 61
CP 709 CATGTTGATACATCCGCCCCAGGTGCGGACATGAAGCTTCTGCC 660

RESULT 8
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag= a
FT /function= multiple cloning site
FT primer_bind 187..204
FT /*tag= b
PN EP-285123-A.
PD 05-MAY-1988.
PF 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Koivula A, Bamford J, Reinikainen T;
DR WPI: 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

Query Match 1.8%; Score 41; DB 1; Length 204;
Best Local Similarity 10.7%; Pred. No. 9.08e-08;
Matches 12; Conservative 60; Mismatches 39; Indels 1; Gaps 1;
DB 75 yycdchvcgymrttthyrmbvnyrdynrsdaawycyrrsvk-ydcynachhd 133
OY 10 TCCGCACTGCTCACTCCCGCGCAGTGAGTGTGCACAGCCCGCTGTGCGTCTG 69
DB 134 dhvyvbbvynvnhnncnccbnhvhvbnhnrwayvrharrdv 185
OY 70 GTTCCCTTAGTCCGAGCGCTCGCCCACTGCAGATTCTTCCCGTGACAGAC 121

RESULT 9
ID N81164 standard; DNA; 204 BP.
AC N81164;
DT 08-NOV-1990 (first entry)
DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69

```

FT      /*tag= a
FT      /function= multiple cloning site
FT      primer_bind      187..204
FT      /*tag= b
PN      EP-285123-A.
PD      05-MAY-1988.
PF      30-MAR-1988; 105163.
PR      03-APR-1987; US-034819.
PA      (SUZO) SUOMEN SOKERI OY.
PI      Lehtovaara P, Knowles J, Kolvula A, Bamford J, Reinikainen T;
DR      WPI; 88-279927/40.
PT      Introducing random point mutations into nucleic acids -
PT      by prepn of single stranded template, annealing a primer, elongation,
PT      misincorporation, completion of molecules and screening.
PS      Disclosure: p; English.
CC      Random point mutations were introduced into the alpha fragment of
CC      E.coli beta-galactosidase. The wild type sequence was obtained as a
CC      single stranded template and an oligonucleotide was hybridised to
CC      it to generate a popn of DNA molecules which terminate at all
CC      possible nucleotide positions within a specified region. The
CC      variable 3' ends generated in this way are used as primers for
CC      reverse transcriptase. Nucleotides are misincorporated by the
CC      transcriptase and the molecules are completed to forms that can be
CC      amplified and then expressed in a suitable host-vector system.
CC      The sequence covers all 176 diff base substitutions, most of which
CC      occurred singularly in any given mutant.
CC      See also P80575.
SQ      Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;

```

Query Match	1.8%;	Score 41;	DB 1;	Length 204;
Best Local Similarity	11.9%;	Pred. No. 9.08e-08;		
Matches	12;	Conservative	52;	Mismatches 37;
			Indels	0;
			Gaps	0;

```
Db      76 ydcchvgccgymrttllhyrimdbnvryrdynrsdaawycocyrrsvkyddccynachdh 135
       |: :| |::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
CP     1435 CCAAGGACGCCTGTATTTTAAAGCTTTTGATGCATTTTACCCAAATGCCCTCAGAAA 1376

Db     136 yyvbbbvynvnhnncnccbnhvchnvbnnhrwayvr 176
       :::::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
CP     1375 TACTGTACCATTTCAGAGTCGCATCATATAATACCAGAA 1335
```

```

RESULT      10
ID          070470 standard; DNA; 114 BP.
AC          070470;
DT          10-APR-1995 (first entry)
DE          Generic DNA sequence to generate a random TSAR peptide library.
KW          TSAR: totally synthetic affinity reagent; synthetic; binding domain;
KW          effector domain; concatenated heterofunctional protein; linker;
KW          direct; rapid; detection; screening; treatment; generic; ss.
OS          Synthetic.
FH          Key
FT          misc_feature      Location/Qualifiers
FT          misc_feature      55..60
FT          /*tag= a
FT          /note= "encoded by 2 (see comments)"
PD          W09418318-A.
PD          18-AUG-1994.
PE          01-FEB-1994; U00977.
PR          01-FEB-1993; US-013416.
PR          30-DEC-1993; US-176500.
PR          31-JAN-1994; US-189331.
PA          (UYNC-) UNIV NORTH CAROLINA.
PI          Fowlkes DM, Kay BK;
PI          WPI: 94-279739/34.
DR          P-PSDB; R58378.
PT          Identifying proteins or peptide(s) which bind a ligand - by
PT          screening a recombinant vector library expressing fusion proteins
PT          comprising a binding domain and an effector domain
PS          Disclosure; Page 36; 255pp; English.
CC          Q70470 is a generic DNA sequence used to generate random TSAR (Totally
CC          Synthetic Affinity Reagents) peptides. This generic formula can also be
CC          represented as follows: X(NNB)4(CAC)(NNB)8Z(NNB)6(CAC)(NNB)8
CC          -(CAC)2(NNB)Y. X and Y are flanking restriction sites (X is not the same

```

as Y) that are not specified further. The peptides generated by this and other generic sequences (Q70471-73) have invariant histidine residues incorporated into variant sequences. TSARs are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains. The TSARs or compsns. comprising a TSAR binding domain can be used in vivo to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody production. The TSARs are easily characterised and have designed activity allowing direct and rapid detection in a screening process.

Query Match	1.58;	Score 34;	DB 12;	Length 114;
Best Local Similarity	9.88;	Pred. No. 2.35e-04;		
Matches	11;	Conservative	29;	Mismatches 72;
			Indels	0;
			Gaps	0;

[illegible]

RESULT	11
ID T31452	standard; cDNA; 3018 BP.

DT	15-SEP-1996	(first entry)
DE	Human Ndr serine/threonine kinase cDNA.	
KW	Serine/threonine kinase; Ndr; nuclear Dbp2-related kinase;	
KW	calmodulin; agonist; antagonist; therapy; ss.	
OS	Homo sapiens.	
FH	key	location/Qualifiers
FT	CDS	596..1993

```

FT      CDS      596..1993
PN      /*tag= a
PD      WO9619579-A1.
PE      27-JUN-1996.
PR      20-DEC-1995; E05052.
PA      22-DEC-1994; EP-810746.
PI      (CIBA ) CIBA GEIGY AG.
PI      Hemmings BA, Millward TA;
DR      WPI: 96-309593/31.
DR      P-PSDB: R98366.
PT      Nuclear Dbf2-related protein kinase - used in the identification of
PT      agonists and antagonists, for treatment of disorders associated with
PT      calmodulin response anomalies
PS      Claim 7; Page 33-37; 48pp; English.
CC      A cDNA clone (T31452) codes for a novel human nuclear Dbf2-related
CC      protein kinase, Ndr (R98366), that binds to calmodulin in a
CC      calcium-dependent manner. The clone was obt. from HeLa cDNA
CC      by PCR amplification using primers (see also T31453-54) based on
CC      conserved regions of C. elegans cmlb8 and Drosophila Ndr (see
CC      also R98365). The Ndr gene was localised to human chromosome
CC      6, between 6p21.2 and 6p21.31, a region that contains major
CC      histocompatibility class I genes. The isolated cDNA can be used
CC      to produce recombinant Ndr in transformed host cells, e.g.
CC      Escherichia coli JM109 or mammalian COS-1 cells, or as a probe
CC      to identify related sequences.
SQ      Sequence 3018 BP; 908 A; 640 C; 672 G; 798 T;

```

Query Match	1.5%;	Score 35;	DB 21;	Length 3018;
Best Local Similarity	74.0%;	Pred. No. 7.88e-05;		
Matches	54; Conservative	0; Mismatches	19; Indels	0; Gaps 0;

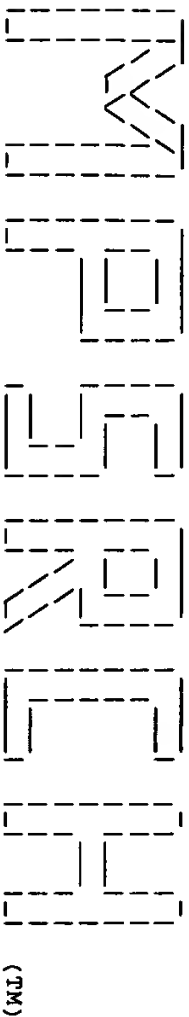
Db 449 tttctctgcttctgttacatatatcctgtttgtccctcaganaacttcacccaatttac 508
||| | | |||||| | || ||||| | | ||||||| | | |||||||
CP 1419 TTTTAAAGTCCTTTTGATGCATTTTACCAATTTGCCCTCCAGAATACTGTACCATTTC 1360

OS	Synthetic.		Location/Qualifiers
FH	Key		misc_feature 55..60
FT	misc_feature		/tag- a
FT	note= "encoded by z"		
PN	WO9418318-A.		
PD	18-AUG-1994.		
PF	01-FEB-1994; U00977.		
PR	01-FEB-1993; US-013416.		
PR	30-DEC-1993; US-176500.		
PR	31-JAN-1994; US-189331.		
PA	(UYNC-) UNIV NORTH CAROLINA.		
PI	Fowlkes DM, Kay BK;		
DR	WPI: 94-279739/34.		
PT	P-PDSB; R58383.		
PT	Identifying proteins or peptide(s) which bind a ligand - by screening a recombinant vector library expressing fusion proteins comprising a binding domain and an effector domain		
PS	Disclosure; page 36; 255pp; English.		
CC	Q70472 is a generic DNA sequence used to generate random TSAR (Totally Synthetic Affinity Reagents) peptides. This generic formula can also be represented as follows: X(NNB)1(CAC)(NNB)11(CAC)(NNB)(CAC)(NNB)22(NNB)6-(CAC)(NNB)5(CAC)2(NNB)4. x and y are flanking restriction sites		
CC	(x is not the same as y) that are not specified further. The peptides generated by this and other generic sequences (Q70470-73) have invariant histidine residues incorporated into variant sequences. TSARS are concatenated heterofunctional proteins or peptides, comprising at least two functional regions - a binding domain with affinity for a ligand and a second effector peptide portion that is chemically or biologically active. They may further comprise a linker peptide between the 2 domains.		
CC	The TSARS or compsns. comprising a TSAR binding domain can be used in vivo to deliver a chemically or biologically active moiety, eg. metal ion, radioisotope, peptide, toxin or enzyme, to the specific target or on the cell. They can also replace the function of macromolecules, eg. monoclonal or polyclonal antibodies and therefore circumvent the need for complex methods of hybridoma formation or in vivo antibody production. The TSARS are easily characterised and have designed activity allowing direct and rapid detection in a screening process.		
CC	Sequence 114 BP; 6 A; 12 C; 0 G; 0 T;		
SO			
	Query Match	1.4%; Score 31; DB 12; Length 114;	
	Best Local Similarity 10.1%; Pred. No. 5.82e-03;		
	Matches 11; Conservative 28; Mismatches 70; Indels 0; Gaps 0;		
Dd	3 bcacnbnbnbnbnbnbnbnbnbnbnbnbnbnbnbcacnbnbcacnbnbnbnnnnnn 62 : : : : : : : : : : : : : :		
Qy	42 GCACAGCACC CGCTCTGTGGCTCGCTTGCTTCCCTTAGTCCGAGCGCTGCCCATCA 101 : : : : : : : : : : : : : : : : :		
Dd	63 bnbnnbnbnbnbnbnbcacnbnbnbnbnbnbnbnbcaccacnbnbnbnbnbn 111 : : : : : : : : : : : : : : : : :		
Qy	102 GATTCCTTTCCCGTGCGAGACATGGCCCTTGCGACACGACACCGCGCGT 150 : : : : : : : : : : : : : : : : :		

RESULT	ID	15
AC	Q70465	standard; DNA; 114 BP.
DT	05-APR-1995	(first entry)
DE	Generic DNA sequence to generate a random TSAR peptide library.	
KW	TSAR: totally synthetic affinity reagent; synthetic; binding domain;	
KW	effector domain; concatenated heterofunctional protein; linker;	
KW	direct; rapid; detection; screening; treatment; generic; ss.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	misc_feature	55..60
FT	/*tag- a	
FT	/note= "this sequence represents 'Z'; 2 can be a	
FT	sequence of 6, 9 or 12 nucleotides (see	
FT	comments)"	
PN	W09418318-A.	
PD	18-AUG-1994.	
PF	01-FEB-1994;	U00977.
PR	01-FEB-1993;	US-013416.
PR	30-DEC-1993;	US-176500.

[illegible]

Search completed: Mon Feb 10 22:55:11 1997
Job time : 231 secs.



Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nn n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 22:55:36 1997; MasPar time 1092.42 Seconds
916.033 Million cell updates/sec
Tabular output not generated.

Title: >US-08-674-311-1
Description: (1-2269) from US08674311.seq (1 of 2)
Perfect Score: 2269
N.A. Sequence: 1 GAATTCGCTCCGACACTGCT.....AAAAAAGCGGAATTC 2269
Comp: CTTAGCGGAGGCGGTGACGA.....TTTTTTTTCGCCCTTAAG

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

EST-STS

1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

Database:

EST-STS-TWO

100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:EST129
130:EST130 131:EST131 132:EST132 133:EST133 134:EST134
135:EST135 136:EST136 137:EST137 138:EST138 139:EST139
140:EST140 141:EST141 142:EST142 143:EST143 144:EST144
145:EST145 146:EST146 147:EST147 148:EST148 149:EST149
150:EST150 151:EST151 152:EST152 153:EST153 154:EST154
155:EST155 156:EST156 157:EST157 158:EST158 159:EST159
160:EST160 161:EST161 162:EST162 163:EST163 164:EST164
165:EST165 166:EST166 167:EST167 168:EST168 169:EST169
170:EST170 171:EST171 172:EST172

173:enSTS1 174:enSTS2

Statistics: Mean 12.257; Variance 2.969; scale 4.129

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No.
1	67	3.0	168	50	HUMNK726	Human keratinocyte CD	2.35e-57
2	49	2.2	400	134	G17388	human STS SHGC-14428	8.74e-32
3	49	2.2	400	173	HS388293	human STS SHGC-14428	8.74e-32
4	41	1.8	473	60	N45983	YY57b12.s1 Homo sapie	2.97e-21
5	33	1.5	344	45	HSIMB133	H.sapiens partial mRN	1.53e-11
6	27	1.2	284	48	HUM231B05B	Human aorta CDNA 5'-e	4.76e-05
7	28	1.2	285	51	MMGEG113	M.musculus DLA-113 mR	4.53e-06
8	27	1.2	431	86	R50764	Y970C04.s1 Homo sapie	4.76e-05
9	28	1.2	452	55	N30827	Yw70a07.s1 Homo sapie	4.53e-06
10	27	1.2	490	168	HSW03788	za61g02.r1 Soares fet	4.76e-05
11	27	1.2	490	147	W03788	za61g02.r1 Soares fet	4.76e-05
12	27	1.2	492	132	G08836	human STS CHLC.ATA21A	4.76e-05
13	27	1.2	565	60	N45982	YY57b12.r1 Homo sapie	4.76e-05
14	25	1.1	144	87	R56208	Y990a03.s1 Homo sapie	4.29e-03
15	25	1.1	186	166	HS933319	zb57e06.s1 Homo sapie	4.29e-03
16	26	1.1	228	147	W04550	zb91e04.s1 Soares par	4.68e-04
17	26	1.1	228	163	HS550335	zb91e04.s1 Soares par	4.68e-04
18	26	1.1	254	132	G08783	human STS CHLC.ATA5H0	4.68e-04
19	26	1.1	263	161	HS362315	zb70d01.s1 Homo sapie	4.68e-04
20	26	1.1	263	143	N93362	zb70d01.s1 Homo sapie	4.68e-04
21	25	1.1	290	49	HUM337G10B	Human aorta CDNA 5'-e	4.29e-03
22	25	1.1	305	48	HUM252H12B	Human aorta CDNA 5'-e	4.29e-03
23	25	1.1	306	132	G09779	human STS CHLC.ATA24E	4.29e-03
24	25	1.1	306	55	N29618	yw67d03.s1 Homo sapie	4.29e-03
25	26	1.1	317	7	H03560	yj37h04.r1 Homo sapie	4.68e-04
26	25	1.1	329	57	N36600	yx88a03.r1 Homo sapie	4.29e-03
27	26	1.1	330	84	R45024	yg36h09.s1 Homo sapie	4.68e-04
28	26	1.1	332	62	N52043	yz08h05.s1 Homo sapie	4.68e-04
29	25	1.1	342	61	N47304	yy74h07.r1 Homo sapie	4.29e-03
30	26	1.1	347	135	HS217ZC3	H.sapiens (DIS436) D	4.68e-04
31	26	1.1	359	46	HUM020E05A	Human fetal brain CDN	4.68e-04
32	26	1.1	380	174	HSB334ZA5	H.sapiens (D2S2301) D	4.68e-04
33	26	1.1	380	136	HSB334ZA5	H.sapiens (D2S2301) D	4.68e-04
34	26	1.1	396	161	HS413323	za73c06.r1 Soares fet	4.68e-04
35	26	1.1	396	146	W01413	za73c06.r1 Soares fet	4.68e-04
36	25	1.1	402	46	HUM065F03A	Human fetal brain CDN	4.29e-03
37	25	1.1	421	12	H19304	ym44a06.s1 Homo sapie	4.29e-03
38	25	1.1	421	62	N51355	yz16a05.s1 Homo sapie	4.68e-04
39	25	1.1	439	9	H08801	yl18a03.s1 Homo sapie	4.29e-03
40	25	1.1	452	53	N24083	yx67a07.s1 Homo sapie	4.29e-03
41	25	1.1	471	83	R42451	yg02b02.s1 Homo sapie	4.29e-03
42	25	1.1	472	143	N92751	zb26a07.s1 Homo sapie	4.29e-03
43	25	1.1	472	165	HS751318	zb26a07.s1 Homo sapie	4.29e-03
44	26	1.1	533	73	R06864	yf11e10.s1 Homo sapie	4.68e-04
45	26	1.1	685	153	W22325	65E9 Human retina CDN	4.68e-04

ALIGNMENTS

RESULT	1	
LOCUS	HUMNK726	168 bp mRNA
DEFINITION	Human keratinocyte CDNA, clone 726.	EST
ACCESSION	D29573	01-DEC-1994
NID	g599506	
KEYWORDS	EST(expressed sequence tag).	
SOURCE	Homo sapiens Epidermis Keratinocyte CDNA to mRNA.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.	
	1 (bases 1 to 168)	

AUTHORS Konishi,K., Morishima,Y.-I., Ueda,E., Nonomura,K., Kibe,S.,
Yamanishi,K. and Yasuno,H.
JOURNAL Unpublished (1994)
REFERENCE 2 (sites)
AUTHORS Konishi,K., Morishima,Y., Ueda,E., Kibe,Y., Nonomura,K.,
Yamanishi,K. and Yasuno,H.
TITLE Cataloging of the genes expressed in human keratinocytes: analysis
of 607 randomly isolated cDNA sequences
JOURNAL Biochem. Biophys. Res. Commun. 202 (2), 976-983 (1994)
MEDLINE 94324994
COMMENT Submitted (11-Mar-1994) to DDBJ by:
Kiyofumi Yamanishi
Department of Dermatology
Kyoto Prefectural University of Medicine
Kamigyo-ku
Kyoto, Kyoto, 602
Japan
Phone: 75-251-5587
Fax : 75-251-5587
PROJECT='human_epidermal_keratinocyte_ESTs'.
location/Qualifiers
1..168
/organism="Homo sapiens"
/cell_line="primary culture"
/cell_type="Keratinocyte"
/sequenced_mol="cDNA to mRNA"
/tissue_type="Epidermis"
26 c 37 g 62 t
BASE COUNT 43 a 26 c 37 g 62 t
ORIGIN
Query Match 3.0%; Score 67; DB 50; Length 168;
Best Local Similarity 93.3%; Pred. No. 2.35e-57;
Matches 83; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Db 19 agattggaataa-t-gtggacaacagcctgagatgacagagaataattagaaggaagaact 77
OY 153 AGATTGGAATAATTGTTGGAACAGGCGCTGATGATCCAGAAATTTTAGAAGGAAGAACTG 212
Db 78 ttaaatatgtgatgac-ccattggcaag 105
OY 213 AAAAATATGTGATCTCTCCATTGGCAAG 241
RESULT 2
LOCUS G17388 400 bp DNA STS 05-MAR-1996
DEFINITION human STS SHGC-14428 clone pg-7320.
ACCESSION G17388
NID 91214814
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human plasmid clones, generated from a lymphoblastoid cell line
from a human male. Localized to human chromosome 9 by analysis on
the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell
Institute for Medical Research, Camden, NJ 08103.
Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Eukaryotes; Metazoa; Chordata;
Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
Catarrhini; Hominoidea; Homo.
1 (bases 1 to 400)
Myers,R.M.
Unpublished (1996)
REFERENCE Contact: Richard M. Myers
AUTHORS Stanford Human Genome Center (SHGC)
JOURNAL Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myerseshgc.stanford.edu
Primer A: TGTGATCTCCATTGGCA
Primer B: TCACAAATCAAGCCAGCAAG
STS size: 90

PCR Profile:
Initial incubation: 94 degrees C for 90 seconds
Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600
Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul
Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3
Chromosome 9.
location/Qualifiers
1..400
/organism="Homo sapiens"
/note="human"
61..150
/map="9"
61..80
/map="9"
primer_bind complement(131..150)
/map="9"
BASE COUNT 95 a 74 c 83 g 101 t 47 others
ORIGIN
Query Match 2.2%; Score 49; DB 134; Length 400;
Best Local Similarity 98.2%; Pred. No. 8.74e-32;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 28 atccagaaatttagaagaaga-ctgaaaaatagtgtgatactccattggcaag 82
OY 186 ATCCAGAAATTTTAGAAGGAAGAACTGAAATAATATGTGATCTCCATTGGCAAG 241
RESULT 3
ID HS388293 standard; DNA; STS; 400 BP.
AC G17388;
DT 07-MAR-1996 (Rel. 47, Created)
DE 07-MAR-1996 (Rel. 47, Last updated, Version 1)
KW human STS SHGC-14428 clone pg-7320.
OS primer; sequence tagged site; STS sequence.
OC Homo sapiens (human)
OC Eukaryota; Animalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Hominoidea.
RN [1]
RP 1-400
RA Myers R.M.;
RT ;
RL Unpublished.
CC Contact: Richard M. Myers Stanford Human Genome Center (SHGC)
CC Stanford University School of Medicine Department of Genetics,
CC M-344, Stanford, CA 94305, USA Tel: 4157259687 Fax: 4157259689
CC Email: myerseshgc.stanford.edu Primer A: TGTGATCTCCATTGGCA
CC Primer B: TCACAAATCAAGCCAGCAAG STS size: 90 PCR Profile: Initial
CC Incubation: 94 degrees C for 90 seconds Denaturation: 94 degrees C
CC for 15 seconds Annealing: 62 degrees C for 23 seconds
CC Polymerization: 72 degrees C for 30 seconds PCR Cycles: 30 Thermal
CC Cycler: Perkin Elmer 9600 Protocol: Template: 25 ng Primer: each 1
CC uM dNTPs: each 200 uM Taq Polymerase: 0.05 units/ul Total Vol: 10
CC ul Buffer: MgCl2: 2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.3
CC Chromosome 9. NCBI g1: 1214814
FH key location/Qualifiers
FH FH

FT source 1..400
/organism="Homo sapiens"
/note="human"
FT STS 61..150
/map="9"
FT primer_bind 61..80
/map="9"
FT primer_bind complement(131..150)
/map="9"
SQ Sequence 400 BP; 95 A; 74 C; 83 G; 101 T; 47 other;
Query Match 2.2%; Score 49; DB 173; Length 400;
Best Local Similarity 98.2%; Pred. No. 8.74e-32;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Db 28 atccagaatttagaaggaaga-ctgaataatgtgatactccattggcaag 82
|||||
OY 186 ATCCAGAAATTTTAGAGAGAGAACTGAAATATGTGATCTCATTTGGCAAG 241
RESULT 4
LOCUS N45983 473 bp mRNA EST 14-FEB-1996
DEFINITION yy57b12.s1 Homo sapiens CDNA clone 277631 3'.
ACCESSION N45983
NID g1187149
KEYWORDS EST.
SOURCE human clone=277631 primer=ml3 -40 forward library=Soares multiple
sclerosis 2NbHSP vector-pt7T3D (Pharmacia) with a modified
polylinker V_TYPE: phagemid host=DH10B (ampicillin resistant)
Rsitel-Not I Rsitel2-Eco RI 46 year old male. 1st strand CDNA was
primed with a Not I - oligo(dT) primer
[5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT-3'],
double-stranded CDNA was size selected, ligated to Eco RI adapters
(Pharmacia), digested with Not I and cloned into the Not I and Eco
RI sites of a modified pT7 vector (Pharmacia). Library went
through one round of normalization to a Cot = 5. Library
constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4
multiple sclerosis lesions from one patient was kindly provided by
Dr. Kevin G. Becker (NINDS/NIH).
ORGANISM Homo sapiens
Eukaryota; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 473)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 399
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES
source location/Qualifiers
1..473
/organism="Homo sapiens"
/clone="277631"
/note="human"
<1..>473
MRNA 187 a 69 c 80 g 137 t
BASE COUNT
ORIGIN

Query Match 1.8%; Score 41; DB 60; Length 473;
Best Local Similarity 79.2%; Pred. No. 2.97e-21;
Matches 76; Conservative 0; Mismatches 17; Indels 3; Gaps 2;
Db 2 gaatattcaaacctcttattacgtgaagaataagccaattt--gttagattac-a 58
|||||
OY 905 GAATGCTCAGAAACCCTCCATTAACCTGAAGAAATATGCCCCAGTTTCTGTTTATTACCA 964
Db 59 agacattaaataacatgctgcccagagaaaga 94
|||||
OY 965 AGACATTAAAGTAGCATGGCTGCCAGAGAAAAGA 1000
RESULT 5
LOCUS H5IMB133 344 bp RNA EST 04-DEC-1995
DEFINITION H.sapiens partial mRNA; expressed sequence tag (clone lmb-est
133).
ACCESSION X93833
NID g1103448
KEYWORDS EST; expressed sequence tag.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 344)
Moschonas,N.K., Argyrokastritis,A. and Vrontou,S.
Unpublished
2 (bases 1 to 344)
Moschonas,N.K.
Direct Submission
Submitted (23-NOV-1995) Moschonas N.K., Institute of Molecular
Biology and Biotechnology, FORTH, P.O. Box 1527, Heraklion 711 10,
Crete, Greece
COMMENT K562-CDNA library normalization: Puzyrev et al., Molecular Biology
(Molecular Biology and Biotechnology), 29:58-61(1995).
both strands read (asymmetric PCR sequencing)
no significant homology to GenBank entries.
FEATURES
source location/Qualifiers
1..344
/organism="Homo sapiens"
/cell_line="K562 (erythro-leukemic)"
/dev_stage="embryonal/fetal with respect to globin
expression"
/clone_lib="normalized K562 CDNA library"
/clone="lmb-est 133"
<1..>344
MRNA 96 a 89 c 72 g 87 t
BASE COUNT
ORIGIN
Query Match 1.5%; Score 33; DB 45; Length 344;
Best Local Similarity 77.0%; Pred. No. 1.53e-11;
Matches 47; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
Db 253 atccataaagaagaagtccttctcttaccacattactcctcagtcaggaagt 312
|||||
CP 1353 ATCAATAATAACCAAGAGAGTCCATTCTCTCTCATTCTTTGGCAGTATAGGAAGT 1294
Db 313 a 313
|
CP 1293 A 1293
RESULT 6
LOCUS HUM231B05B 284 bp mRNA EST 28-AUG-1995
DEFINITION Human aorta CDNA 5'-end GEN-231B05.
ACCESSION D62007
NID g965783
KEYWORDS EST(expressed sequence tag); Human aorta;
similar to none(May 29,1995).
SOURCE Homo sapiens (Library: Clontech human aorta polyA+ mRNA (#6572))
CDNA to mRNA.
ORGANISM Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;

Vertebrata; Sarcopharygii; Mammalia; Eutheria; Primates;
Cetartihni; Homiidae; Homo.
1 (bases 1 to 284)
Fujiwara,T., Hirano,H., Katagiri,T., Kawai,A., Kuga,Y., Nagata,M.,
Okuno,S., Ozaki,K., Shimizu,F., Shimada,Y., Shinomiya,H.,
Takaichi,A., Takeda,S., Watanabe,T., Takahashi,E.-I., Hirai,Y.,
Maekawa,H., Shin,S. and Nakamura,Y.
Unpublished(1995)
Submitted (30-May-1995) to DDBJ by:
Tsutomu Fujiwara
Otsuka GEN Research Institute
Otsuka Pharmaceutical Co.,Ltd
463-10 kagasuno Kawauchi-cho
Tokushima, Tokushima
771-01
Japan
Phone: 0886-65-2888
Fax : 0886-37-1035.
Location/Qualifiers
1..284
/organism="Homo sapiens"
/sequenced_mol="cDNA to mRNA"
/clone_lib="Clontech human aorta polyA+ mRNA (#6572)"
BASE COUNT 87 a 27 c 42 g 104 t 24 others
ORIGIN
Query Match 1.2%; Score 27; DB 48; Length 284;
Best Local Similarity 66.7%; Pred. No. 4.76e-05;
Matches 34; Conservative 5; Mismatches 12; Indels 0; Gaps 0;
Db 6 atgtaaatyytgtaaaatagataamataataactgctgyatgt 56
|||||||:|||||||:|||||||
Oy 1734 ATGTAATATACAAAAAAGTAGAAGAAATATATATACCTGTATTGT 1784
RESULT 7 MMGEGL13 285 bp RNA EST 14-SEP-1995
LOCUS
DEFINITION M.musculus DLA-113 mRNA.
ACCESSION X71637
NID 9297509
KEYWORDS expressed sequence tag; transcribed sequence fragment.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Sarcopharygii; Mammalia; Eutheria; Rodentia;
Sciurognathi; Myomorpha; Muridae; Murinae; Mus.
1 (bases 1 to 285)
REFERENCE
AUTHORS del Mazo,J.
TITLE Direct Submission
JOURNAL Submitted (27-APR-1993) to the EMBL/GenBank/DBJ databases. J. Del
Mazo, Centro de Investigaciones Biologicas, C.S.I.C., Velazquez
144, Madrid 28006, SPAIN
2 (bases 1 to 285)
Lopez-Alanon,D.M. and del Mazo,J.
REFERENCE
AUTHORS Lopez-Alanon,D.M. and del Mazo,J.
TITLE Cloning and characterization of genes expressed during
gametogenesis of female and male mice
JOURNAL J. Reprod. Fertil. 103 (2), 323-329 (1995)
MEDLINE 95341612
FEATURES
location/Qualifiers
1..285
/organism="Mus musculus"
/strain="Swiss"
/tissue_type="ovary"
/dev_stage="fetal"
/clone="DLA-113"
BASE COUNT 88 a 62 c 65 g 70 t
ORIGIN
Query Match 1.2%; Score 28; DB 51; Length 285;
Best Local Similarity 80.4%; Pred. No. 4.53e-06;
Matches 37; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 212 tgcccttttaaatgaagcttctgtcagaaaaaaaaa 257
||||||| ||||| ||||| ||||| ||||| |||||
Oy 2215 TGCCCTTTAATAAAGTGAGAGCTTGCTTTTAAAAA 2260
RESULT 8
LOCUS R50764 431 bp mRNA EST 18-MAY-1995
DEFINITION YG70C04.s1 Homo sapiens cDNA clone 38721 3', similar to gb:M58583
CEREBELLIN PRECURSOR (HUMAN);.
ACCESSION R50764
NID R50764
KEYWORDS 9812666
SOURCE EST.
human clone=38721 library=Soares infant brain INIB vector=Lafmid BA
host=DH10B (ampicillin resistant) primer=Promega -21ml3 Rsite1-Not
I Rsite2-Hind III Whole brain from a 73 days post natal female. 1st
strand cDNA was primed with a Not I - oligo(dt) primer (5'
AACTGGAAGATTGCGCGCGCAGGAATTTTCTTTTCTTTT 3'); double-stranded
cDNA was ligated to Hind III adaptors (Pharmacia), digested with
Not I and directionally cloned into the Not I and Hind III sites of
the Lafmid BA vector. Library went through one round of
normalization. Library constructed by Bento Soares and M.Fatima
Bonaldo.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopharygii; Choanata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Cetartihni; Homiidae; Homo.
1 (bases 1 to 431)
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
REFERENCE
AUTHORS The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT
TITLE
JOURNAL
COMMENT
GDB: G00-411-262
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 298
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
FEATURES
location/Qualifiers
1..431
/organism="Homo sapiens"
/clone="38721"
/note="human"
BASE COUNT 130 a 90 c 77 g 128 t 6 others
ORIGIN
Query Match 1.2%; Score 27; DB 86; Length 431;
Best Local Similarity 84.6%; Pred. No. 4.76e-05;
Matches 33; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Db 8 ttttttttttttaaaagtcaccacttcgactttat 46
||||||| ||||| ||||| ||||| ||||| |||||
Cp 2260 TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT TTTT 2222
RESULT 9
LOCUS N30827 452 bp mRNA EST 05-JAN-1996
DEFINITION YW70A07.s1 Homo sapiens cDNA clone 257556 3', similar to gb:K00558
TUBULIN ALPHA-1 CHAIN (HUMAN);.
ACCESSION N30827
NID 91149347
KEYWORDS EST.
SOURCE human clone=257556 primer=m13 -40 forward library=Soares placenta

/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGAGAGATATATTAAGATCTTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo." /clone="297074" /sex="male" /dev_stage="20 week-post conception fetus" /lab_host="DH10B (ampicillin resistant)" <1..>490

MRNA
BASE COUNT 149 a 91 c 64 g 186 t

Query Match 1.2%; Score 27; DB 147; Length 490;
Best Local Similarity 84.6%; Pred. No. 4.76e-05;
Matches 33; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Db 319 ctttttttttttttaagaatgaatcatcattttat 357
|||||
CP 2261 ctttttttttttttAAAAAGCAAGCTTCACCTTATT 2223

RESULT 12
LOCUS G08836 492 bp DNA STS 08-AUG-1995
DEFINITION human STS CHLC.ATA21A09.P15975 clone ATA21A09.
ACCESSION G08836
NID 9939386
KEYWORDS STS sequence; primer; sequence tagged site.
SOURCE human vector-pbCPI host-E.coli dut+ung+ (DH10B) Marker selected genomic DNA prepared from xy individual of French nationality.
ORGANISM Homo Sapiens
Eukaryota; Eukaryotes; Metazoa; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 492)
Murray,J., Sheffield,V, Weber,J.L., Duyk,G. and Bueltow,K.H.
Cooperative Human Linkage Center
Unpublished (1995)
Synonyms: ATA21A09, CHLC.ATA21A09.#T15719
Contact: Dr. Jeffrey C. Murray
UofI
The University of Iowa
Department of Pediatrics, Iowa City, IA 52242, USA
Tel: (319) 356-3508
Fax: (319) 356-3347
Email: jeff-murray@uiowa.edu

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
PRIMER A: ATGCAATCCAGAGCTTCAAC
PRIMER B: ATGATATGTGCTGGGAGA
STS size: 337
PCR Profile:

denature: 30 seconds at 94 degrees C
annealing: 75 seconds at 55 degrees C
extension: 15 seconds at 72 degrees C
PCR cycles: 27
extension: 6 minutes at 72 degrees C
Protocol:
Template: 30ng genomic DNA
Primer: each 1.5 pmole
dNTPs: each 200 uM
Taq Polymerase: 0.3 units
Total Vol: 10 uL

Buffer:
MgCl2: 1.5mM
KCl: 50mM
Tris: 10mM

FEATURES
source
PH: 8.3.
Location/Qualifiers
1..492
/organism="Homo Sapiens"
/note="human"
STS
primer_bind 71..407
primer_bind 71..90
complement(388..407)
BASE COUNT 231 a 76 c 75 g 103 t 7 others
ORIGIN

Query Match 1.2%; Score 27; DB 132; Length 492;
Best Local Similarity 73.6%; Pred. No. 4.76e-05;
Matches 39; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

Db 270 taataataataataataaagccctgaanacnttgcaaccttaca 322
|||||
QY 1519 TAATAATGATAGTTATATATTAATAATTTGAACAATCTGAATCCCTTGCAA 1571

RESULT 13
LOCUS N45982 565 bp MRNA EST 14-FEB-1996
DEFINITION YY57B12.r1 Homo sapiens cDNA clone 277631 5'.
ACCESSION N45982
NID 91187148
KEYWORDS EST.
SOURCE human clone-277631 primer-T7 library-Soares multiple sclerosis 2NbhMSP vector-pT73D (Pharmacia) with a modified polylinker V_TYPE: phagemid host-DH10B (ampicillin resistant) Rsite1-Not I Rsite2-Eco RI 46 year old male. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCGCATTTTCTTTTCTTTT-3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from 4 multiple sclerosis lesions from one patient was kindly provided by Dr. Kevin G. Becker (NINDS/NIH).
Homo sapiens
Eukaryota; Metazoa; Bilateria; Coelomata; Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 565)
Hiller,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 409
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.

FEATURES
source
1..565
/organism="Homo sapiens"
/clone="277631"
/note="human"

MRNA
BASE COUNT 194 a 80 c 100 g 188 t 3 others
ORIGIN

THIS PAGE LEFT BLANK


```

source
1..2269
/clone="18-11 CDNA"
/map="9p21"
/chromosome="9"
/organism="Homo sapiens"
/sex="male"
/cell_line="primary culture"
/cell_type="fibroblast"
/tissue_type="epidermis"
122..973
/gene="MTAP"
/codon_start=1
/product="methylthadenosine phosphorylase"
/db_xref="PID:g847724"
/translation="MASGTTTAVKIGIGTGGLDPELIEGRTEKYVDTPGKPSDA
LILGIKNVDCILLRHGRQHTIMPSKYNQANIMALKKEECTHIVITACGSLREI
QPGDIVIIDQFIDRTTMRPOSFYDGSCARGVCHIPMAEPFCPTREVLITPAKLG
LRCHSGTMTVITIEGPRSSRAESEFMRTWGADVIMNTIVDEVVLAKKAGICVASIAMA
TDYDCWKEHEEAVSVDRVLTKLTKENANKAKSLLTITPQIGSTEWSETLHNLNMAQF
SVLLPRH"

BASE COUNT      725 a      407 c      490 g      647 t
ORIGIN

```

Query Match	100.0%;	Score 849;	DB 56;	Length 2269;
Best Local Similarity 100.0%;	Pred. No. 0.00e+00;	Mismatches 0;	Indels 0;	Gaps 0;
Matches 849;	Conservative	0;	Mismatches 0;	Indels 0;
Db 122 atgacctctgacacacacacacacacgcccgtgaagattggaataaattggtggaacagcctg	181			
QY 122 ATGACCTCTGACACACACACACACCGCCGTGAAGATTGGAATAATTGGTGAACAGCCTG	181			
Db 182 gatgaccgaagaatttagaaggaagaactgaaataatgtgatactccatttggcaag	241			
QY 182 GATGATCCAGAAATTTTAGAAGGAAGAACTGAAATAATGTGATTAATCCATTGGCAAG	241			
Db 242 ccacatcgtatgccttaatttggggagaataaaaaatgttatgcatcctccttgcgaag	301			
QY 242 CCATCTGATGCCCTTAATTTTGGGGAGAATAAAAAATGTGATTGCATCCTCCTTGCAAG	301			
Db 302 catggaaagcagcacacacatcatatgccttcaaaagttcaactaccagggcgaaatctgggct	361			
QY 302 CATGGAAGGACGACACACCATCATGCGCTTCAAAAGTCAACTACCAGGCGAATCTGGGCT	361			
Db 362 ttgaaggaagaggctgtacacatgtcatagtaccacagcttctggtcccttgaggag	421			
QY 362 TTGAAGGAAGAGGGCTGTACACATGTCTATAGTGACCACAGCTTGAGGCTCCTTGAGGAG	421			
Db 422 gagattcagcccgcgatatgttcattatgtacagttcatattgacaggaaccatataga	481			
QY 422 GAGATTCAAGCCCGCGATATTGTTCATTATTGATCAGTTTCATTGACAGGACCACATATGAGA	481			
Db 482 cctcagtccttctatgatatgtaagtcattcttgcgcagagaggtgtgcataatccaatg	541			
QY 482 CCTCAGTCTCTTATGATGGAAGTCATTTCTGTGCCAGAGAGTGTGCCATATTCCAATG	541			
Db 542 gctgagccgcttctgcccccaaaacgagagaggttcttatagagactgtctaagaagctaga	601			
QY 542 GCTGAGCCGTTTGGCCCCCAAAACAGAGAGAGGTCTTATAGAGACTGCTAAGAGCTAGGA	601			
Db 602 ctccggtgccaactcaaaaggggacaatggtcaacaatcgagagacctglttagctcccg	661			
QY 602 CTCGGGTGCCACTCAAAAGGGGACATGCTCACAAATCGAGGAGCCTCGTTTAGCTCCGG	661			
Db 662 gcaagaaagcttcatgttccgcacactggggggcgatgttatcaacatgacacagttcca	721			
QY 662 GCAGAAAGCTTCATGTTCCTCGACCTGGGGGGGATGTTATCAACATGACCACAGTTCCA	721			
Db 722 gaggtgttctctctaaggaaggtctggaaattgttcaagcaagtatcgcattgacagacagat	781			
QY 722 GAGGTGTTCTTGTCTAAGGAGGCTGGCAATTGTGTACGCAAGTATCGCCATGCGCACAGAT	781			
Db 782 tatgactgctggaagagcacaggaagcaagtttcggtgacccgggtcttaagaacctg	841			

QY	782	TATGACTGCTGGAGGAGCAGCAGGAGGACAGTTTCGGTGGACCGGGGTCTTAAAGACCCCTG	841
Db	842	aaagaaaaacgctaataaaggccaaaagcttactgctcactaccataccatagataggtcc	901
QY	842	AAAGAAAACGCTAATAAAGCCAAAAGCTTACTGCTCTACTACCATACCTCAGATAGGGTCC	901
Db	902	acagaatgtcagaaaacccctccataacctgaagaatatggccagtttctgtttatca	961
QY	902	ACAGAAATGGTCAGAAACCTCCATAA CCTGAAGAATATGGCCAGTTTCTGTTTATTA	961
Db	962	ccaagacat	970
QY	962	CCAAGACAT	970

RESULT	2			
LOCUS	SPAC16C9	14071 bp	DNA	PLN
DEFINITION	S.pombe chromosome I	cosmid c16C9.		22-OCT-1995
ACCESSION	Z54366			
NID	g1019812			
KEYWORDS	MAY32 protein homologue; NAM7 protein homologue; NOT4 homologue; serine/threonine protein kinase; zinc finger protein.			
SOURCE	lission yeast.			
ORGANISM	Schizosaccharomyces pombe			
	Eukaryotae; mitochondrial eukaryotes; Eumycota; Ascomycotina; Euraeascomycetes; Schizosaccharomycetales;			
	Schizosaccharomycetaceae; Schizosaccharomyces.			
	1 (bases 1 to 14071)			
REFERENCE	Badcock,K. and Churcher,C.M.			
AUTHORS	Unpublished			
JOURNAL	2 (bases 1 to 14071)			
REFERENCE	Barrell,B.G., Rajandream,M.A. and Walsh,S.V.			
AUTHORS	Direct Submission			
TITLE	Submitted (12-OCT-1995)			
JOURNAL	to the EMBL/Genbank/DBJ databases.			

Centre, Hinxton Hall, Hinxton, Cambridge CB10 1RQ E-mail: barrellesanger.ac.uk

Details of yeast sequencing at the Sanger Centre are available on the World Wide Web.
(URL, <http://www.sanger.ac.uk/yeastpub/swv/home.html>)
Protein coding regions (CDS) have been predicted with the help of computer analysis using the GeneFinder program in Pombase (an ACEDB database) with additional predictions for the branch-acceptor sites supplied by the program Sp3splice. CAUTION: It is possible that for any individual CDS we may have underestimated or overestimated the number of introns/exons or we may not have chosen the correct splice donor/acceptor sites.
CDS are numbered using the following system eg SPAC5H10.01c. SP (S pombe), A (chromosome 1), c5H10 (cosmid name), .01 (first CDS), c (complementary strand).
The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous.
The length in codons is given for each CDS.
IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions.
Cosmid c16c9 is overlapped by cosmid c4g8 at the 5' end and by cosmid c2g11 at the 3' end.

```
FEATURES
  source
    location/Qualifiers
    1..14071
    /organism="Schizosaccharomyces pombe"
    /strain="972h-"
    /chromosome="I"
    /clone="cosmid c16C9"
    /map="TL"
    1..267
    /note="overlap with 4G8"
    complement(96..101)
    /note="splice donor sequence, gtaagt"
    complement(join(>100..179,217..256))
```


Dp 1196 gaatgactgcacatgttcatatactggctccccagccttcgtgatataattgcttcgcac 1255
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 719 GAAGTGTGTCATGTTGATACATCCGCCCGGCTGGGACATGAAGCTTCTGCC 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Dp 1256 ggcgtacaataatgcaggaccctccat 1280
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Cp 659 GGGAGCTAAAACGAGGCTCCCTCGAT 635

RESULT 3
LOCUS CELB0228 38946 bp DNA INV 30-MAR-1995
DEFINITION Caenorhabditis elegans cosmid B0228.
ACCESSION U23168
MID 9726361
KEYWORDS
SOURCE ORGANISM
Caenorhabditis elegans strain=Bristol N2.
Eukaryotae; mitochondrial eukaryotes; eukaryote crown group;
Metazoa/Eumycota group; Metazoa; Eumetazoa; Bilateria;
Pseudocoelomata; Nematoda; Secernentea; Rhabdita; Rhabditida;
Rhabditina; Rhabditoidea; Rhabditidae; Caenorhabditis.
REFERENCE
AUTHORS
1. (bases 1 to 38946)
Wilson,R., Ainscough,R., Anderson,K., Baynes,C., Berks,M.,
Bonfield,J., Burton,J., Connell,M., Copsey,T., Cooper,J.,
Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Laister,N.,
O'Callaghan,M., Parsons,J., Percy,C., Rifken,L., Roopra,A.,
Saunders,D., Showkeen,R., Smaiden,N., Smith,A., Sonhammer,E.,
Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,
Wilkinson-Sproat,J. and Wohlman,P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C
elegans
JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE
AUTHORS
2. (bases 1 to 38946)
Leimbach,D.
TITLE The sequence of C. elegans cosmid B0228
JOURNAL Unpublished (1994)
REFERENCE
AUTHORS
3. (bases 1 to 38946)
Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (21-MAR-1995) Robert Waterston
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RO, England
e-mail: rtw@nematoe.wustl.edu and jesus@sanger.ac.uk

NEIGHBORING COSMID INFORMATION:

The 5' cosmid is B0495, 8001 bp overlap; 3' cosmid is C06A8, 750 bp
overlap. Actual start of this cosmid is at base position 4495 of
CELB0228; actual end is at 6105 of CELC06A8

NOTES:

Coding sequences below are predicted from computer analysis, using
the program GeneFinder (P. Green and L. Hillier, ms in
preparation).

FEATURES
source Location/Qualifiers
1..38946
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
join(56..502,1135..1583,1735..2017,2219..2342,2390..2842,

CDS

2988..3076,3122..3241,3287..3355,3404..3472,3519..5054,
/gene="B0228.4"
/map="iti"
/note="coded for by C. elegans cDNA yk26d12.5; coded for
by C. elegans cDNA yk21b9.5; coded for by C. elegans cDNA
yk16f1.5; coded for by C. elegans cDNA yk32h8.5; coded for
by C. elegans cDNA yk16f1.3; coded for by C. elegans cDNA
yk32h8.3; coded for by C. elegans cDNA yk26d12.3; coded
for by C. elegans cDNA yk21b9.3"
/codon_start=1
/db_xref="pid:q726370"
/translation="MADYEIRDLTGSSGSSNNOKISNNGSESAHENGTSNKKTESEDD
DSTCSSLNKLNSNRHSEPMNHLVSGFEVAKDLISLTIPLPQTIVRTQSIDIDTPDR
SRNDKLTLSEFYSSCKNLLTTITIDPLDLVAAGVQSTQAQLVAQIGMWCYCCRC
CLCKCIANGRGVGRGKQKQIQEPTIADDDGRDLPEYPGSIHIPKYNNKLKSRSS
VPYVSPSLVSSADYKRLSMETLADSNVESDEVEPLTPPLVGRKLDYIDKEGNPSC
CRERGRKPGFRPLRGSGSPPPDPDSTKVAVYTSAEVKKIKHKKQKLEKEKKK
MDRKSSSGGGFTWMEGSGSNSSQNLAEVIDARSEKTAQOREQELLQREBRSS
GGRTHSEERHRCPOPSLMVNTNDEDDYATIDRYRNSNRNEMSPASPRNVHDET
SGPLSTMNESAFAFGRAHLQYRPAOKGAAGPTRGAPTSVNOEDLADTIDRLSTE
PSPVSRRLPKSAISSVQOKQPIKTADGGQLIGSVYTWQNSVDTTTDDRKADF
LRGBSSRLSPQSEKRNKQRIQROQSSGPTNRETEIEYEERKOGPPVTRVEGKL
KMERVAGADLTPOSCISSAWTVDTVNTKIKSTIGKSLISLEIKDGQSKYITLI
ENGFERKMERASLDVDPFVNKDYLAESVKSLSLDLREDSSESALTHIEVEIVEDT
NLTQYVIGERAGADVAEOLRLHYEOTADKTPDSPIPLEKVIYVDELQKRIELED
PERKIDHLIKDGRHEGEGALKRIRAFETEESTKEPTVIMREPCAAHAFDAVAKD
DTSNTYVKIAVPLVHTITFLLKSKMRQOKAGYEMEOEGREGFEDTLTRIKRYET
EESEKOHAVIIOHEEVKVAVMTKTQREVEABEGQYEMSQEGIHLRGETTRKRGHL
DSESEERFMERAEAGGOYAMNEMGERELIGKERFSKGRHRESESESMASNSGPT
LVDLYKKESSSIEATFEATFANNHSPVIAEVRPKLKENTTGCTISNOKATSANAE
TKHVHNTKESGSKFRELGEQAMMLCGFENQSSKEEVTGSNOQKNEIKVAFAGSAE
TENTTISTITFDADSAFAVEGSSKSANSTATGRFKEMSEENASNMVYLQKSESSSN
LSGAETTKMKHIOROSSEARFSESKQVAESCAVMIKNTGVGSGTSSVAEATG"
join(6960..14594,15067..18774,19220..21034)
/gene="B0228.3"
/name="iti"
/codon_start=1
/evidence="not_experimental
/db_xref="pid:g726364"
/translation="MEESVSEVYCECKRRSSSIDRASLKLKLAQSEVCTGFNTGSEST
RRTILQKTKVESMCLTKTSASSTSIDMSDCQATVSGTISTEIASHREIVAAFG
ISTQSLREKLEVEIIDWENTMNSOKEDLSANLQALTSAPVNDIPEDLGRIVAPA
BODAYSDLINREARESVFAPOLRASADELTRIVALGNSQQLQAAFMELITSVSR
DLRTIAPSNITANTEIFVYAAEKMASVSGWKRRESKRESKFTSREIIOGFWR
REDEERYERKLDKRMVSFKHSLSVEAASSTSENVLGMKADQKSEPTISOKLSPKEI
VSEAYGVSEKLDQFQVMEKMDWSNIEGEKHTALSNIRTLAPEQVCDGILGK
KAPKQODESVGTQIOEIRATVAMSVRSLLTITSSNSFSKNITSEKAVFSNLISV
ITSHNLSTIGTSETTASISYQDIPEMLAASKLWISNNEKLLKKEIREPIQTVESF
WNTINDQEKETVSSDLQGFENVLEKMDMSQISLPAQEHAIISKVRSLAPASASELSNL
RELVSSEKFTVSSDLQGFENVLEKMDMSQISLPAQEHAIISKVRSLAPASASELSNL
GKLRAPEQDEETDLKIQAOQARVYLVNVESSMNTISSETFSRIPENKATITNLV
GIITQDLSIKGASENCTIEMDYTEAKELEBASKLVNKLNSVLKTEIRESGEVYQ
GFMTVAQAEFRISSESDHLTIHEKMDMSKIDIVEEYSSTISNNYQALATSSQSGS
STRDVQAQAEFRISSESDHLTIHEKMDMSKIDIVEEYSSTISNNYQALATSSQSGS
ILGKRLPRAPBEETSKEREINTAKCILNVISLSTTSDASLDMISEQALATPSN
IIGVAMSNLNTISSQSSISSEFGNNVELSEARVQESNKNOLKVRSESEELHG
IMSTASKEKEVAVIYKREKLETVHQAKMTLAIQMATQSVNSELIGSEGNLASIKSTLPL
TREVISAAGFISNEVIOKMLEVLITGEMWSTISLPOSEYESIQNVRAALAEPSFNDISI
LGKINAPGPOTETTSHELLIEKNTAAVANKVAVESVISKESSLAKLPADAKAITNL
AGVLAKSDVSSLCSTSEETFGFDQRTFRNNQANISLGAVRSETTLQRLREPIENQVGF
WSTASQDEKASFIAXOKLDTMDYKVIYAIOISQITDGDSSASISSETFKNIDKAA
REVAVAEFGIANESVQKALEVLANKVEMSDISLPELHQNISNVARLADSSYNDOSII
GKLNAEPQASAVDQISEHOTLEITANIKSAVSYSINDSVLSRSEDEKAVISNTA
NLISDCLASMSSTNSSEFQDRIEPOANVAVLCTPASQVFTNLQEPISOEYQGF
STNSOQKASLVIRKLVNVDAMKMATETIAYQSLNSNTSEENLSEKSPFQSR
ESIMASFGVSEHVOKTEILKSLSTVLPVNEYEATVONIKALSEPNEBCSILG
KLNAPDQASAMVDIMNEOKRLIVYNSIQASIAESVINDSSLKLELTVKSVIKKE
IIVSDTLTASCSGSDENLOESNTQEKADIFLKPNSQMLIEHLREPIEKVQGFSE
TASNEKQEMFLKEVETIHAMLOTFSASLVEVYORDFMATVQSLAALRSTQITPRE
ILPCAQFISTNEQETRGLEDVMSHIDIPICARNDLNLVNRVYNADAPNVFNGLIS
RPEQAGETISNEQETRGLEDVMSHIDIPICARNDLNLVNRVYNADAPNVFNGLIS
ENLGDILISQAVOLAOPNMSDETEKSFEPINRLINRVVETSEESIOSEFWKTSOLAES

ASSTISEKLSMLOSEFVVSAAQVSTSLTLDYRRKIFNONSEIIFGDLTRDVVKAAS
VSDETLNTLFSOLEOSDWSQIKLSRQKSLISANIKSLATSLNLDLTLGOLMPREADSE
OSEASLERSSVEMENTEFKIQOILFDYTVRKFSGSDEERASIKINLIGLTVNLKAIQI
VASFERLPDSLDSTASITPVEFKIOEHRQPTQVQGWFSKSRHPNVEYVNRKIEV
LKSILNCYAAVECHKSLERSLEKPSQODECATRGVTEVQVNAFAVDEQSYSKILVL
GSMQNFEPFAVKETISKNLKILNLPVDTLLVSKYKESHVSKTVEQDELVOFMENK
SALNSEIEMSQNSVNCIEEETVVLPLNVLASLANVATIPVTESSRQSDSENILNFI
OHESESASATVRDENILNRHYLOTPAIOKASSEVRAVIEIKTSQSFHLVHTMLEE
IKTIOLEMEVLKRVSMNOSTAEELLQAFONSEOOMETFORIVQNHMEETKLEEI
RNLRIINTAI PREFESTVGLKAPDDQHESADSYISSTRAEAVINLMAADNAISTSS
TLMSESDERLLHKTIILKTMAVCNLTTPASRSETETMTQGFYQKVEQCOTEQELERP
KINCEMKACVETTCGTETENMKNPDDAEITTEFVALTGLTSLCDLVNTAKSNIOVDSKYD
YRRPARKNAETTFCTESNADAFSALOGEAVLTSSGIMTWSNSDAKTITVSDKLISI
SKTEMSMAYASENAVNODHSLKSDISEYEMNIPDSRKEILOKNYSIDRSNSTVEMOG
AIASEDLAIQYSAPRVDNVQOTMRQSOQDILLFGSGFGELEPLPOEEDVETARQSR
VYRSSQVAPSEESIQTKEALRRSESESNARKTFVDRRRENVASIRKASVERETNM
EAVAKADOSIPVETLOKSCKQESIFESVASESKDLNVCGSWTAKPPIGAVSLQTK
VEKEVTSAMTVASASVECEVGLSEKRSQSDATGSMVRKASIEEYEREEFGEVSTSTE
KTLERQDQIESWIKISIPORLEVEESAEGNDEVITGGVMSLEAPLEQOEEETEKLTLM
KRSASEARSLKATKESIEBEADEFSKTEADQDEILTVOKELVATGILINAASREINAS
SKMEYSKPYSEDIELSLTESRQSNQSGQFETKEEIEVGLMNTGTGENASVLPKPK
PPIDTASMKAKAKAKONSIEMTGSLOKSASAMGIVSQVTEGANSKFGIAQAVET
LTASAQSGTCRDISVNSIGNIVASETVQFSNKETGIGFASNLAPBESAETEFTGK
ISNLQTSLNMKAASDVGTVEKIQAPGDNYGDVSLQKVASSDAITKMAQASRDSA
ISYDFROBERTVSAEKSDLNFKSTNSEOKLFESEKEESGIFIRSHXEYETOKTL
RHRASRSBASRTVAPTNOEVQMPDKVEDVSAEGSLSIGIVRESSQSEVMQHAER
TSELTKLSMNEEVAGVRAVASETTFNETFPRYQOQDVAEVQGTGAMGRIEARPQGEAER
TOKLRLTSLVERSAKASAMASQVTOQKREDSLASEYSVRDTLLKSSSVSHVATE
QMEHLYMRSKSEHHISEKLOERLIEREFSQOEFISENOGVHTHNDVIDNGDALIC
WKSSETEQKSLDAKQVTEELAGTTLDSVRLGQDEEKIAEHVIGASEVLNISETR
ADMENSRQDAFSDQATVASDILIELRSLSTVHEFGSEASATTFGVGLVTKPKEEY
GRSFSETRKLSQFSDITAISETTDMOSELTLPECDQSVTLISHRNOOKQDKLQAT
VENTADQVTLLEKREVRQOSADVKKKFEVEDIEEVLPMIRMAELISMKASTSVEIQ
AENKLSKBEAQASOKYLKTANLEINKINYNATHEVMASATTALECKTSGQDVASIRL
RDKSREREKQAOENQMNLSLTNAEMENLDEDSVTIAQSVQDSMAFSVKASATVN
LNSDTISKKAOPEIGIOKSLISQSNVEKTYQGFSSATIGNELVLRGMDLLEIETLV
EINREQANAGRIREFGSETGGIYLVRALPKVKETSHTVATVATSFQOIFSTMSAG
DEISEANVELTIPSTVGAIEIKSTVARSMSMTFSTSHASEYTASTVADYARDIIVASAS
TAARKKAIPTERTSOQLKEVGADGIEILSMEGIEETFDATQTLVDLVRKSSLOQIE
SSEEVERNOLYLEMPEEKGLVHLINYNOKETCERNFAVISCSDLYVSRKSEIAPT
SVSKYTERKYLRETRVATVIESGDVRFNANVLNHRYSLSKPTLAQEVTLREVRIAPAO
LFTISAGDIESDVVWSTHOLEKTFVREEDITTIIGARKGENIKKKFEAGDEKVKLVE
LIGSNLPNEICDISWKIPRLGKVKKLETEEEFEDECLFLMSNQLQADNF"

join(21201..25576,25621..25699,25774..25874,25950..26055,
26445..26576,26630..26780,26826..26900,26948..27083,
27129..27228,27467..27586)

/gene="B0228.2"

/map="iii"

/note="coded for by C. elegans cDNA yk19c1.5; coded for by
C. elegans cDNA yk45a10.5; coded for by C. elegans cDNA
yk21e5.5; coded for by C. elegans cDNA yk21e5.3; coded for
by C. elegans cDNA yk35c1.3; coded for by C. elegans cDNA
yk45a10.3"

/codon_start=1

/db_xref="PID:g726363"

/translation="MRTKESSEDIQVGGIAYNPPEENLISAIKLDEKASGGCYELTTK
AAGDEYKMSISLSTQAVEKVKGMQIKVAKDEIRVLETTESVANMENYEIPEN
FKIGIAIKTANNAPFTMRALHSCDLQFYKFMNMNKKDDFSLIKITISNGIESKTL
SCIAEHIETARNPEFNRPAAEFSDIQTYVDYKNIEGASDLNTEFALNEMKALSTMLSR
ODDYSRVEHIVKDKNRGANLKRFRIESSSEKQTFSAFIDEITEKEIEERTIDLVRQGG
QFKLSTDAENEITLHRDISKPIITHYDILHLITLSNAPHQILSTGSASNELHTIS
TQLSKPAEWNTTELLIVDKNVEQPTVTRVECEBCEVENLHPYRRPDIJFDLEIWHI
ARNGKGFERRCKASGDEKEIFEQIEKISRGTKDDIIDKFIIVNGQEPSTFTTIQTSV
SASVSQDSMRGEKEAIKVLVNSNGKINVEKNIATEYRETISEOFRNDEDFKAD
LTVKDRAGSGYELSTNASQSSSVSNELICPRPSQLSIEKTFITAOITIPAILSCK

...
Note: remainder of annotations omitted.

Query Match	3.7%;	Score 31;	DB 25;	Length 38946;
Best Local Similarity	65.0%;	Pred. No. 1.52e-06;		
Matches	67;	Conservative	0;	Mismatches 36;
				Indels 0;
				Gaps 0;

Db 33432 catcactgtgttccatatgtgtatctacgcaactgtgactggtatccagtaaatat 33491

```

CP      251  CATCAGATGGCTTGCCAAATGAGATATCCACATATTTTTCAGTCTCTCCCTCTAAATTT 192
Db      33492 ttgatcttccaaccacttctccaataatccaacctgac 33534
          ||||| ||| | ||| ||||| ||| | ||||| ||| |
CP      191  CTGATCATCCAGGCCCTGTCCACCATTATTCCTTCTTAC 149
          ||||| ||| | ||| ||||| ||| | ||||| ||| |

RESULT  4
LOCUS   RRCYTBCL 5550 bp DNA BCT 15-FEB-1991
DEFINITION R. rubrum cytochrome bcl-complex genes (petA, petB, petC).
ACCESSION X55387
NID      946382
KEYWORDS cytochrome b; cytochrome bc-1 complex; cytochrome c1; petA gene;
          petB gene; petC gene; Rieske Fe-S protein.
SOURCE   Rhodospirillum rubrum.
          Rhodospirillum rubrum.
ORGANISM Eubacteria; Proteobacteria; alpha subdivision; Rhodospirillum.
          1 (bases 1 to 5500)
REFERENCE 1 (bases 1 to 5500)
AUTHORS  Majewski, C.
TITLE     Direct Submission
JOURNAL   Submitted (05-NOV-1990) to the EMBL/GenBank/DBJ databases. C.
          Majewski, RUHR UNIVERSITAET BOCHUM, LEHRSTUHL F BIOCHEMIE DER
          PFLANZEN, RUHR UNIVERSITAET, POSTFACH 102148, 4630 BOCHUM 1
          2 (bases 1 to 5550)
REFERENCE 2 (bases 1 to 5550)
AUTHORS  Majewski, C. and Trebst, A.
TITLE     The pet genes of Rhodospirillum rubrum: cloning and sequencing of
          the genes for the cytochrome bcl-complex
JOURNAL   Mol. Gen. Genet. 224 (3), 373-382 (1990)
MEDLINE   91094774
FEATURES
          source
          Location/Qualifiers
            1..5550
            /organism="Rhodospirillum rubrum"
            /strain="Frl"
            /clone_lib="Partially digested Sau3A fragments of genomic
            Rs. rubrum DNA in EMBL3."
            /clone="EMBL-pet 11.1, EMBL-pet 4.1"
            <3..>770
            /gene="ORF1"
            3..770
            /gene="ORF1"
            /codon_start=1
            /transl_table=11
            /db_xref="PID:946383"
            /translation="MLDAPAGRWPPDPDAGEGALRRLDGEKDLVIDRALLRWAERG
            KGPYPRAERTARWRLIETALGWTAPTEPLGGDALACLGPAVGEAMRPCADAGLR
            AARPGRTRCWPGSPPLGGTGPPTNPPRSRAMNRYSDRPSGVLRPACGMAOGLAE
            HOGGDDGVERAHAGNHRNPDAQVALVDLGNARALSAOOHVYGLKVSFGMKRRL
            RCQDQAGFVGEPRDEIRPGRMTRSGA FEIVHSGAA"
            prim_transcript <3..>770
            /gene="ORF1"
            1250..1253
            RBS
            prim_transcript <1261..>1812
            /gene="petA"
            <1261..>1812
            /gene="petA"
            1261..1812
            /gene="petA"
            /codon_start=1
            /transl_table=11
            /product="Rieske Fes protein"
            /db_xref="PID:946384"
            /translation="MAEAHTASTPGESSRRDFLIYGTAVGAVGVALAVMPFIDFM
            NPAADTLASTEDVSAIAEGQALTVWRGKPVFVRHRTQKEIVVARVDPASLRDP
            QTEARVQAQMLVMVGVCVTHLGCIPLGQAKADPKGDFDGFWCFCGSHYDSAGRIRK
            GPAPLNLPPVPAFTDDTVLLIG"
            1811..1814
            RBS
            prim_transcript <1826..>3043
            /gene="petB"
            1826..3043
            /gene="petB"
            /codon_start=1
            /transl_table=11

```



```

/product="cytochrome b"
/db_xref="pid:g46385"
/translation="MYTPPRNNKALKWEDERLPVLTVAHKELVYYPAPRLNYSWFN
GSLAGIAMITMIATGIFLAMSTAHVDHAFSEVERIMRDVNGWLMRYMHANGASMEF
IVVYVHMFGRGLYGSYKPREVLWMLGLVILLMTATAFMGYLLPWGQMSFGCATVIT
NLFSAPVVGDDIVTLTMGGSVDNPTLNREFSLAYLEPMLFEPAVFEILHMLALPKS
NNPLGIDAKGPEDTIPFHRYTVKDAFGIFLMVCEFEVFAPEFGEPDNYIAPN
MVTPTHVPEWYFLPFYALIRAVPDKLGGVAMFGAILILEVLPWLDTSKVSATERP
LVKGFWEFLADCLLLGYLGAMPAAEERYVYITQLATIYFYLHFLVITPLVGMEEKPKP
LPVSISSPVTQA"
<1826..>3043
/gene="petB"
RBS 3056..3059
CDS 3070..3888

```

```

mRNA      <3070..>3888
           /gene="petC"
sig_peptide 3070..3142
           /gene="petC"
prim_transcript <3070..>3888
           /note="cytochrome c1"
           /gene="petC"
mat_peptide 3143..3885
           /gene="petC"
           /product="cytochrome c1"
RBS      4053..4056
prim_transcript <4069..>4785
           /gene="ORF2"
CDS      4069..4785

```

```

/conserved_table=1
/translation="MSEAYRQPVLGIGSGVYDIDGLEGRMQTVESPFQGVSDQIL
RGTLDGLEMAFLPRHGRGVLA PSDVNTNRANIDALKRAGYTEILSVSAVGSLEADLP
GTEVLAQDAEIDRTFAREKSEFROGSGRPGCHGSGOALGRSRRRGSGRSGHSPGR
HLTVAGGAVALDPGRKOSLSAMGLPRHRNDQHARQAAGRSPRSRDLHLRHGRHFRLLA
PRSPRQRORGGGSRRAAKRR"
/sequence=">4785
/sequence="ORF2"
RBS
prim_transcript
4983..4986
<4997..>5550
/sequence="ORF3"
mRNA
<4997..>5550
/sequence="ORF3"
4997..>5550
DS

```

BASE COUNT	ORIGIN
926 a	1816 c 1767 g 1041 t

	Query Match	3.28;	Score 27;	DB 22;	Length 5550;
	Best Local Similarity	59.2%;	Pred. No. 1.61e-03;		
Matches	87; Conservative	0;	Mismatches 60;	Indels 0;	Gaps 0;
Dd	4267	ccttcgcatgtgaactacgcgcgcgaatatcgacagccctgaagcggcgcgctgaccgag	4326		
Oy	326	ccttcaaggctcaactaccaggccgaacatcttgacctttgaaaggaaagggctgtcacacat	385		

Db	4327	atcttgtcgtgtctccgcgcgtgcgattcctctgscgaggaacctgcgcgcgcgcacattcgtc	4386
QY	386	GTATAGTAGACCAACAGCTGTGGCTCCTTGAGSAGAGGAGATTCAGCCGGCGATATTGTC	445
Db	4387	atcgccgatacgtttcatcgcgcgcacc	4413
QY	446	ATTATTGATCAGTTCATTGACAGGACC	472

RESULT	5	HUMNFKB34	3205 bp	mRNA	PRI	15-OCT-1990
LOCUS		Human factor KBf1	mRNA,	complete cds.		
DEFINITION		M55643 M37492				
ACCESSION	NID	g189179				
KEYWORDS		factor KBf1.				
SOURCE		Human, cDNA to mRNA.				
ORGANISM		Homo sapiens				
REFERENCE		Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homidae. 1 (bases 1 to 3205)				
AUTHORS	TITLE	Kleran,M., Blank,V., Logeat,F., Vandekerckhove,J.S., Lottspeich,F le Bail,O., Urban,M.B., Kourilsky,P., Baeuerle,P.A. and Israel,A. The DNA binding subunit of NF-kappa-B is identical to factor KBf1 and homologous to the rel oncogene product				
JOURNAL		Cell 62, 1007-1018 (1990)				
MEDLINE		90367112				
FEATURES		Location/Qualifiers				
source		1..3205				

```

/sequenced_mol="cDNA to mRNA"
182..3091
CDS

```

BASE COUNT	874 a	784 c	840 g	707 t
ORIGIN				

Query Match	2.8%	Score 24	DB 60	Length 3205
Best Local Similarity	78.6%	Pred. No. 2.03e-01		
Matches	33	Conservative	0	Mismatches 9; Indels 0; Gaps 0;
Db 1287	ccaatttcgagatgttcgscggtgtgtagtgcgcgag	1328		
Cp 168	ccaattattccaattcttcacggcggtggtggtgcacagag	127		

RESULT	6			
LOCUS	HUMNFKB	3625 bp	mrna	
DEFINITION	Human nuclear factor kappa-B DNA binding subunit (NF-kappa-B)			20-FEB-1991
ACCESSION	M58603	complete cds.		mrna
NID	g189177			
KEYWORDS	nuclear factor kappa-B DNA binding subunit.			

REFERENCE		HUMAN premyceloid, cDNA to mRNA.
AUTHORS		Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae. 1 (bases 1 to 3625)
TITLE		Meyer, R., Hatada, E.N., Hohmann, H.-P., Haiker, M., Bartsch, C., Roethlisberger, U., Lahm, H.-W., Schlaeger, E.J., van Loon, A.P.G.M. and Scheidereit, C. Cloning of the DNA-binding subunit of human nuclear factor kappa-B: The level of its mRNA is strongly regulated by phorbol ester or tumor necrosis factor alpha Proc. Natl. Acad. Sci. U.S.A. 88, 966-970 (1991) 91126115
JOURNAL MEDLINE FEATURES		
SOURCE		location/Qualifiers 1..3625 /organism="Homo sapiens" /cell_line="HL60" /sequenced_mol="cDNA to mRNA" /tissue_type="premyceloid" /tissue_lib="lambda ZAP" 398..3304 /codon_start=1 /product="nuclear factor kappa-B DNA binding subunit" /db_xref="PID:g189178" /translation="MAEDPYLGRPEQMHLPSTLTHTFENPEVFOPMALPTDGPYL OILEQPKQRGFRFYVCEGSPSHGLPGASSEKKNKSYPQVKICNYGPAKIVQLVTNN GKNILHLAHSLVGHKCDEDGICTAGTCAGVGFANGLIHTKKKVFLEARMTEA CIRGNPGLVLVHPDLAYDAIEGGDRQLGREKELIROALQOTKEMLSVRLMFTA FLPDSTGSTFRLEPVNSDAIYDSKAPNASNLKIVRMDRTAGCTGGEITYLLCDKVQ KDDIQIRFEENEENGWGVEGDFSPIDVAROFAlVEKTPKYKDINTKPASVFQVR RKSDLETSEKPELYYPEIKDKKEVOROKMLPNFSDFGGSGAGAGCGMGSGG GGGGTGTGPGYSFPHYGEPPTYGGITFHPTTKSNAGMKHGMTDESKNDPEGCKSD DKNTVNLFGRVIETTEDOEPESEATVGNGEVTLTYATGKESAGVODNIFLEKAMQL AKRHANALFDYA VTGDVKMLAVQRHLTA VODENDSVLHLAI IHLHSOLVRDLLLEV SGLISDDIINMRNDLYQTPPLHAVITKQEDVEDLLRAGADSLDLRLGNSVLAIAK EGHDKVLSTILKHKAALLLDHPNGDGLNAIHLAMSNSLPCLLLVAGADVNAOEQ KSGRTALHLAVEHDNISLAGCLLEGDAHYDSTTYDGTPLHIAAGRGSTRALAILKA AGADPLVENEBPLYDLDDSWENAGEDGVGCTPLDMATSMQVFDILNGPYEPEFT SDDLAAOGDMKOLAEDVKQLYLKLEIPDPKNWATLAOKLGLILNNAFRLSPAPSK TLMDNVEVSGGTIVRELVEALROMGYTEAIEVIQAASSPVKTTSQASHLPLSPASTROQ IDELRDSDSVCVDIGVETSFRKLSTESTLSGSASLLTLNKMPHYGQEGPLEGKI"
CDS		
BASE COUNT	933 a 937 c 956 g 799 t	
ORIGIN		
Query Match	2.8%; Score 24; DB 60; Length 3625;	
Best Local Similarity	78.6%; Pred. No. 2.03e-01;	
Matches	33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;	
Dn	1500 ccaatttcgagatagttcgccggtgtagtgccggag 1541	
Cp	168 CCAATTATTCCAATCTTCACGCGGTGTGTGCCACAG 127	
RESULT	7	
LOCUS	HSNX1112 9152 bp DNA PRI 18-JUL-1995	
DEFINITION	H.sapiens NFKB1 gene for p50/p105 subunits NF-kappaB (exons 11,12).	
ACCESSION	Z47735	
NID	9902632	
KEYWORDS	NFKB1 gene; p105; p50; transcription factor NF-kappaB.	
SOURCE	human.	
ORGANISM	Homo sapiens	
	Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae. 1 (bases 1 to 9152) Heron, E., Deloukas, P. and van Loon, A.P. The complete exon-intron structure of human NFKB1 encoding the p105 and p50 proteins of transcription factor NF-kappaB Unpublished 2 (bases 1 to 9152) Heron, E. Direct Submission Submitted (12-JAN-1995) to the EMBL/GenBank/DBJ databases. Elke Heron, Vitamins & Fine Chemicals, Biotechnology, F., HOFFMANN-LA	

ROCHE, BASEL, 4002, SWITZERLAND
Location/Qualifiers
1..9152
/organism="Homo sapiens"
/clone="L10, L11, L12, and L13"
/cell_type="B-cell"
/cell_line="Eppstein-Bar transformed B cell line, also
called DANt"
/sex="Male"
<1..160
/note="this is the 3' end of intron 10 that is about 1.7
kbp long"
/number=10
161..299
/gene="NFKB1"
/number=11
join(<161..299,8898..>9041)
/gene="NFKB1"
/note="pid:e122791"
/codon_start=2
/product="p50/p105 subunit(s) of transcription factor
NF-kappaB"
/db_xref="PID:g902633"
/translation="LPLSSKLSIKLILQNPCLSSFGNLTWKLVNQNLSTILK
SKIKKCRGNVNRSSCPIFRIVSAVVVPEDEAEACLVAVEGALEVOVQ"
300..8897
/gene="NFKB1"
/number=11
360..642
/gene="NFKB1"
/rpt_family="Alu"
8898..9041
/gene="NFKB1"
/number=12
9042..>9152
/note="this is the 5' end of intron 12 that is about 1.7
kbp long"
/number=12

BASE COUNT 2884 a 1682 c 1885 g 2701 t
ORIGIN

Query Match 2.8%; Score 24; DB 54; Length 9152;
Best Local Similarity 78.6%; Pred. No. 2.03e-01;
Matches 33; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 8937 ccaatttcgagatagttcggcggtgtagtggtccggag 8978
||||| ||| || | ||||| ||||| ||||| |||
Cp 168 CCAATTATTCATCTTCACGCGGTGTGTGTGTCACAG 127

RESULT 8
LOCUS CELF18C5 29095 bp DNA INV 15-JUN-1995
DEFINITION Caenorhabditis elegans cosmid F18C5.
ACCESSION U29097
NID g861394
KEYWORDS
SOURCE
ORGANISM Caenorhabditis elegans strain-Bristol N2.
Eukaryotae; mitochondrial eukaryotes; Metazoa/Eumycota group;
Metazoa; Eumetazoa; Bilateria; Pseudocoelomata; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditoidea;
Rhabditidae; Caenorhabditis.
1 (bases 1 to 29095)
Wilson,R., Alnscough,R., Anderson,K., Baynes,C., Berks,M.,
Bonfield,J., Burton,J., Connell,M., Copey,T., Cooper,J.,
Coulson,A., Craxton,M., Dear,S., Du,Z., Durbin,R., Favello,A.,
Fulton,L., Gardner,A., Green,P., Hawkins,T., Hillier,L., Jier,M.,
Johnston,L., Jones,M., Kershaw,J., Kirsten,J., Lalster,N.,
Latreille,P., Lightning,J., Lloyd,C., McMurrray,A., Mortimore,B.,
O'Callaghan,M., Parsons,J., Percy,C., Riken,L., Roopra,A.,
Saunders,D., Showkeen,R., Smaildon,N., Smith,A., Sonhammer,E.,
Staden,R., Sulston,J., Thierry-Mieg,J., Thomas,K., Vaudin,M.,
Vaughan,K., Waterston,R., Watson,A., Weinstock,L.,

TITLE
2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans
JOURNAL
Nature 368 (6466), 32-38 (1994)
MEDLINE
94150718
REFERENCE
2 (bases 1 to 29095)
AUTHORS
Fulton, L.
TITLE
The sequence of C. elegans cosmid F18C5
JOURNAL
Unpublished (1995)
REFERENCE
3 (bases 1 to 29095)
AUTHORS
Waterston, R.
TITLE
Direct Submission
JOURNAL
Submitted (13-JUN-1995) Robert Waterston
COMMENT
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: twenematode.wustl.edu and jesus@sanger.ac.uk

NEIGHBORING COSMID INFORMATION:

5' cosmid is C56E6, 200bp overlap; 3' cosmid is C56C10, 200bp overlap. Actual start is at bp 1 of this sequence; actual end is at bp 2639 of CELC56C10

NOTES:

Coding sequences below are predicted from computer analysis, using the program Genefinder (P. Green and L. Hillier, ms in Preparation).

FEATURES

source

CDS

Location/Qualifiers
1..29095
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
join(59..115,607..719,867..1020,1064..1214,2717..3300,3347..3499,3545..3618,3691..4156,4204..4349,5021..5177,5227..6489,6540..6722,6775..6914,6962..7370,7414..7528,7573..7880)
/gene="F18C5.3"
/map="11"
/note="coded for by C. elegans cDNA yk29d2.3; coded for by C. elegans cDNA yk2a12.3; coded for by C. elegans cDNA yk7b7.3; coded for by C. elegans cDNA yk7b7.5; coded for by C. elegans cDNA yk2a12.5; coded for by C. elegans cDNA yk29d2.5; similar to S. cerevisiae hypothetical 287.5 kd protein in PDR region (SP:YBA4_YEAST)"
/codon_start=1
/db_xref="PID:g861397"
/translation="MNADLVKSLRMNVCERLAEIFHAFPRQIGKDEMLVTKNVLNNOYVEETPLISVQRSETNTPAIPFAIVKLIASICRIPSEYOMLTVTINMRTKEQKDVLLKCDALDLILAPLWPGCYDLMETIRNGILKLESLADEPMISYDLVDCQVEOKKRVSGTSLTRKMRILIEFLTKQIEKEVSENKRPAPESTILLERESDFVSESDVYANRLATPLLCYERKKSNEEDLVNMMKSLARIAPLKDPOSYFWRLPKLFSRMEGRLRDSLVTMVEGFVYSMNLDEKTEILHYICELDAMDKTVDPEPHERRYAATTELTKIMNGDVA TNHVDVAMILSSHFHLISTTDLSLRMTAGNNVRAMVOYAGKTLPIAEKRIFLDSYLNPTIIAEPLKHDVAVREELNLTGLVMITGKESQHLAERORAIRKLVESVASEGKNISFNAMNKYLPIMIHPIYLVNRSKAFNAISDESGLIKLIMSKAPWGKCYCSYLESWLARVEKATATKDNEFTDKALVRIYVAIVAEAFHGVSLDEGEKFEISKNOEODEQVSKNEKHTILORINRVILPRLSKSLDSOAVERNAFSAETHATALHLDIQAPIALATYKLLQKLPDAVTGKHUHGVIILKCLNMTHTSYDVRETARKTLVQIVKCGPKYLASVITEISLTMTKGFOVHAIFESVHTLIAMREVTKGELDTAINVIMKCIHQFAEDKENGAVKAECP EAKGNAPEMQMHLGRIVSPAGIOMYIAPFRDVVNEHPSAKAVQVSDLSKFAAGLKDNESLDHSQGLGIFKSLTSDIOKLEVEYTKQKDEKQGRPESCLIPAPORIGAMSKYVIRSDHVAEFYVLLFSSSLKEKFFDLSDSMVSRINLPVKIILDCDFEYKELKISCSMARLSSMIQOLPAIASNSQVSDTLFILLSDSISSIGAGNKPAILQNLQIYKGFNLIIVTIGSDFLDNDKLTLLAYAEADVMDQHQATMSELKALYKRGVHERLGEIMHLSETAIRSPLVNIRSOCRDTLLDIIGSASDAEKSVKEHVEFFLDQLEYEYETGRQSAEMLEALFKNLVAKSLAPLHMLLVYKMGAAAMNEESPKANHIGLALRHLLLES

CDS

VGAQRQETFEVICKWLKAEDENAKAVGIOTAVOLSYVEKGMVTRIDEIVKQVEVL FDDAVENMSETTITVILNGFTRITISNVGASALTKEPDVDFIRSMEECKCEESLGIM LASSTIGQLSHLDVNDIPKELSKDYCFWMSRHLRHEKMDTAIGEQASKNIVCLAKY IDIEDYKTLVGYIAACRVEIKHHVQSLKRINCFKLYAALFVTDGDEENIIVLDAFM SLIVRELKTSDEELSRLTOEVCGLYKRRIGEDAYSNRVATCQKVAAEKITDRKRIR ELAVTAPEDAELKRRKNKKTEVRKRIIDEIKPYRAMKRAAEKRAQENED"
join(8186..8300,8357..8413,8461..9238,9284..9354,9405..9786,10106..10191,10498..10580,10631..10847,10897..11222,11270..11369,11419..11543,11595..11727,11772..11864,11912..12071,12121..12356)
/gene="F18C5.2"
/map="11"
/note="coded for by C. elegans cDNA yk41c3.3; coded for by C. elegans cDNA yk41c3.5; coded for by C. elegans cDNA yk7h8.3; coded for by C. elegans cDNA yk7h8.5; similar to DEAD box helicases"

/codon_start=1
/db_xref="PID:g861396"
/translation="MISDDDDLPTIRPGSVNEELPETEPEDNDELPEPESSDSKPT VTSNKTENQVADEDYDSFDDFVPSQHTASKIPVKNKRAKCTVESDSSSDSDGCDCEETIPACDETOEVPKIKRGYTLRTRASVKNKCDDSDMDGIDEEDVSKRSEDTLNDSF VDPPEMDSVLDNQTLTKGKQFLDDEEFFTDNRVPOIDEATKMKWASMTSPQOALNA LNEFFGHKGFRKQNDVVRNVLGKQDFVLMSTGYGKSVCYQLPSLLNSMTQVSPPL ISLANDQVTVLVSKGIDAVKLDGHSQIEMDQVANNMHRIRFITYSPSEMTVSQGLEL LTRCRKHISLAIDEAHCVSOWGHDNRNSYRHIAEIRNRSDLCNIPMALATATAVR RDDYIANLRKRPLITTTSFDRKNLYISVHSSKDMAEDLGLFMKIDEVKGRHFGPTI IYCOQOMVDVNCVLRIRIGVRSAHYAGLTKNOREKATHDRDKQKEVEENLTMM LRQLELVLTVCGRYQLLKHFDPSYAKPPTMQADCCDRCSTEMLNGNODSSSIVDVT TESKWLFLQVINEMVNGKTGIGKPIEFLRGSSEKEDWRIRKTSQKLEFGIKHIPDKMWK ALASLRIRAGYIGEVRLMOMKFGSCITLSELGERWLLTGKEMKIDATPILLOGKKEKA APSTVPGASRSQSTKSTSEIPTKILGANKIRIEYEPANENEQMLNLRQEVTLGPEKID QLRSLDDIRVGIAMNHEVAPFOIYVNTVLDCFANRPTASASNLMDGMSAOKSRX GKRRYDVCVQFSKENGIGIATNVNANDMIPPELISKQKVLSDAVRKYTEHLISRSTAK EVATARGISEGTVSYLAMAVEKGLPLHLDLNVRKNIAMALNAVRYHLSNVAVLT PWEAMGVDPDENOLKILRAILYEGLDTSENEKRPDIQSMPTSPSTIKTVPSTP SSSLRAPPLKKFKL"

CDS

complement(join(12967..13103,13609..14014,14063..14260,14515..14653,15136..15190,16544..16811,17055..17258,17327..17527,17577..17711,17756..17950,18000..18080,18264..18362))
/gene="F18C5.4"
/map="11"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g861398"

tRNA

CDS

/translation="MATNNSFHYKKEVOFERKNVVVACDVVNGKITVYDVSSOLM GYLFAGADEILNVNENINIKTTADFEIRASAKLPKGVITIEFMRDHLVTEKMPPRRP NTOMVELFTYRGAATGIIHRAVDPKTVTISMVQASNACHYKVDGDIIVKVDI YVLDKRSARKLXYASVNTKTVRLLERSTTDPGRAPRPDPGQASTPAGVKSGNS VMLEPPKPPQGNKNOFQDVLPLPDVLEIMKANKDFYKPKCNPCIKITTAASAPAY PATLPSPAPAEVRPIYIPSPKPKMTKPSHIVREFLLVSGSCGSKSRRLRYV LLIISTSFSSCLICTDRETYETGAAWYRLCAGSCPCPCPPAALTROMTENMKRAM DEKLRLIFYDGLINNLSTIMPAFSPLAATNEKNRPIDVIWSHEDDEEOTIPVKN OGEYAKVFNRSRNQOHAKLIVLAVLAVHFYVILNDFLVKFAAVDQADARYIKPSGIN HRMDVASVRYTQGFIDYKKNLMTVTGAMCASLATITSLTIYSSPRDVFNRAAMAVI RYEDLCAFMEISLELFAIRISSAVIINHQTQHSLRASRISVNLVNDLGNLSCTVPR GELPICSROIIDSIFPVQVIEYLIIICLTGAYIGLAYLIEWCIRHFFPTHEKLTPI SAISRNLVSSRCCSSDSNVLIAA"
18476..18548
/note="lys; codon recognized: AAA; aa: Lys"
complement(join(18753..18876,18926..19369,19411..19477,19837..19870))
/gene="F18C5.5"
/map="11"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g861399"
/translation="MKLLILAVCVIGFVACYEVVPRASALRPBGDVRRSPPLSLAYA PFAKSLRPASTGIRRAPVAVVLPQAOKANRSVSKRAPKPLPMTLPLPLPLPPLPPIV PVPVLASEGLPLPLTAPHTPETETPIPGMPMPALTMPSPFOGLGITTTMKPAEK AMETVQTEESDNDTTSARAHSAPTYSKDLNTRSLRSKFVRGSEKKVTKAEDNADMI VPFH"
join(22223..22486,22538..22683,22771..22846,22947..23316,

CDS

```
23376..23446,23488..23643)
/gene="F18C5.1"
/map="11"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:9861395"
/translation="MAMYNQTAEELETWRCTSDGIFKSONSINAKINFEVFIILIFL
TFYLSFKARVLKHNHYVSKSQIILMITLNLNLIIFLEIRHLVHIFINSEDP
CKIEFHSPECTYDQIYATFSVMSTGLSALTDFRFPALYASTVYVRNSKDSAYMLIT
VSIIVTVIHIRFYGVSRAGVSPCTPQSLNTYQVNNNAIFWIMANCVLTIAY
LLNIYKDKRLIIFYEKLFEENOFRIKSVFDTKTRYSFENVLTKAICSVTSTOF
VFLSESTAALAIIRLEAGMSEVEFHINIQINGGVGNLSIPVLIYLTNQICLQRR
KSIDKMTNHTGTVDSHISSLKTAWET"
complement(join(23918..24100,24145..24215,24262..24386,
26110..26528,26759..26986))
/gene="F18C5.6"
/map="11"
/note="similar to F44F4.7; transposon contained within
intron"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:9861401"
/translation="MSECCARSDVHNVLTSDSMKFNHCFIISIIISFETTTSVRVL
LKONLPTCTRNLFSAIINGIHOCTAVIRLRAFYAIVYASDCAILFOSSOCF
DGNLYYTNLFSSFCFSLFLDLRFSEKPRSSYHNQTLASIVLISQIVLPIGPLYW
VEYDAFTSYVMCTYPPMMSVMKLHEVNIRICVLIVLFFAIFLYIHNKIREKRMV
HNVINNSRYKSYENYLATKSVCIIVFSQILCVGPTSSITSVIFRFRDSIPLEWHLI
ISYLTGLTYSNPLPLIILYQDKQIAKKRRIMQRLQNKNETSFHDFTLSLWGKKT
GNOETLF"
complement(24698..25519)
/gene="F18C5.7"
/map="11"
```

Note: remainder of annotations omitted.

Query Match 2.8%; Score 24; DB 26; Length 29095;
Best Local Similarity 92.9%; Pred. No. 2.03e-01;
Matches 26; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 23819 ttcttgctccagagagtgatgcatatt 23846

Oy 508 TTCTTGTCAGAGAGTGTGCATATT 535

```
RESULT 9
LOCUS CELB0361 42528 bp DNA INV 26-FEB-1994
DEFINITION Caenorhabditis elegans cosmid B0361.
ACCESSION U00031
NID 9458948
KEYWORDS acid phosphatase; transmembrane;
phosphatidylserine decarboxylase.
SOURCE Caenorhabditis elegans strain-Bristol N2.
ORGANISM Caenorhabditis elegans
Eucaryotae; Metazoa; Nematoda; Secernentea; Rhabdilia; Rhabdilita;
Rhabdilitina; Rhabdilitidea; Rhabdilitidae; Caenorhabditis.
1 (bases 1 to 42528)
Wilson, R., Ainscough, R., Anderson, K., Baynes, C., Berks, M.,
Bonfield, J., Burton, J., Connell, M., Copsey, T., Cooper, J.,
Coulson, A., Craxton, M., Dear, S., Du, Z., Durbin, R., Favello, A.,
Fulton, L., Gardner, A., Green, P., Hawkins, T., Hillier, L., Jier, M.,
Johnston, L., Jones, M., Kershaw, J., Kirsten, J., Laister, N.,
Latreille, P., Lightning, J., Lloyd, C., McMurray, A., Mortimore, B.,
O'Callaghan, M., Parsons, J., Percy, C., Rifkin, L., Roopra, A.,
Saunders, D., Showkeen, R., Smaiden, N., Smith, A., Sonhammer, E.,
Staden, R., Sulston, J., Thierry-Mieg, J., Thomas, K., Vaudin, M.,
Vaughan, K., Waterston, R., Watson, A., Weinstein, L.,
Wilkinson-Sproat, J. and Wohlman, P.
2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans [see comments]
JOURNAL Nature 368 (6466), 32-38 (1994)
MEDLINE 94150718
REFERENCE 2 (bases 1 to 42528)
AUTHORS Du, Z.
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

The sequence of C. elegans cosmid B0361
Unpublished (1994)
3 (bases 1 to 42528)
Waterston, R.
Direct Submission
Submitted (26-FEB-1994) Robert Waterston
Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
e-mail: rwenematode.wustl.edu and jesse@sanger.ac.uk

NOTES:

Coding sequences below are predicted from computer analysis, using
the program Genefinder (P. Green and L. Hillier, ms in
preparation).

```
FEATURES
source
location/Qualifiers
1..42528
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
join(670..924,2021..2142,2405..2564,2997..3037,3123..3383,
3657..3942)
/gene="B0361.4"
/map="3"
/note="coded for by C. elegans cDNA CESE79FC"
/codon_start=1
/db_xref="PID:9458952"
/translation="MLSLCFEYILFYNICGASHSSHHHHSHQILYMSTLNFSESV
SNFRLSSDGSGLSLPLTASSNDRRSRLSCSAQMKDDAAYAAQONKDKTGSGSG
RKASSGLVPSLNRRLRIQCFKAKPSIGDAIMKRAASRAEMRTMLSKMNEKQIECG
KOMFELITDAVENADKSEKVSFRGISELLIIFWSTPDPNSVEHMLYVLMASDLISS
PLLMQQLQVSNLMEFIKDVKHCILGANYSLCSLYEEMVLIKEYGIREELRYRRL
SRNSFPSTFRKHPTRASSDKYLKPSIIVSNYFSLKIISSIIFAVNSFFTEKRAHPH
ENFSSSFHPGFVYVDLKSISYCSFSLIYFYEFISIF"
complement(join(4177..4486,4678..5207,5737..5861,
6429..6465))
/gene="B0361.5"
/map="3"
/note="coded for by C. elegans cDNA cm02b7; similar to
phosphatidylserine decarboxylase"
/codon_start=1
/db_xref="PID:9458953"
/translation="MALNPLIFVAFMSWKMLSVSTLLIGGASYVGYLFTPDWREIV
DSKHYISNMKIRVYLSLPENTASRVIGLANOEIPWLRHLGGFARMYDCRMDCV
DPDEFKNYPSFAAFENRKLKESTRPISASPLVSPADGTVLHGEKVEDNKIEYVKGHYD
VDKFLGDVDPKDELPLYOVVILYAPGDYHAFHSPARWANAQCRHVPGLLSVPRPL
LSHVPHFLCLNERVVLNGSWRHGFFSMAVAATNVGDIVDAEPSLRTNIVRKTQKI
MNTETEIHPYVSGERVGESFRLGFTYLVFQAPTIKFAIRAGDPLRYGOSLVADGV"
complement(join(11179..11392,11488..11568,11624..11738,
11844..12220,12301..12463,12515..12604,12653..12749))
/gene="B0361.6"
/map="3"
/note="coded for by C. elegans cDNA CEMS50F; similar to
H. marismortui hypothetical 23.1 kd protein in HMAL3 5'
region"
/codon_start=1
/db_xref="PID:9458954"
/translation="MHPDQKTFEKENDIRKKLFNSTKEERLDWRKMKDEKKRNEEK
IRKEAEAKKAKIEKVDHTPFTISIAVPGOFLNNAQSAELRTYMAQIARAATLYRV
DEIITYDESCRNTDEAVNAYNGTWGNLIPAEITNIEGCFYIAKILEYLECPQRLKRD
LEPIQPLKNGGLNPLDAQHKLKDEKTLRFREGVVLKRSKDGRCICNGLEKEF
EIDSDAVQLPYTRVTVETIKNLTECKLYRGSITSGAIVTRFGLGWGYSVRLMTGLQ
KYLQAKKFDIVAGVSPRGKLASQMDVCLNPKKILLVFGVAGYDAVESEELAEWR
AEDAFDVLIRTSLSNGSRSEERVEENYLSVLAQVQCHLETNAL"
join(13082..13156,13202..13387,13433..13536,14311..14581,
14632..14909,15398..15501,15552..15711,16140..16269,
16313..16402,16447..16620,16732..17111,17157..17421,
17472..17625,17693..17840,17884..18034,18087..18260,
18311..18691,18769..18849,18904..18997,19045..19313)
```

```
/gene="B0361.3"
/map="3"
/note="contains transmembrane regions similar to those in
E. coli tetracycline resistance protein and arabinose
transporter"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g458951"
/translation="MSISRSRYEQFDEMKSENOENNSKKSSERLKLDPDKFEVAVG
AYGKYOIFTYVLOTINFEYSSMYIMSFVOLNLEKCEYKNETIPISETCOIETESS
KAFGNLNGEYCGIAENTLVNTNOKASTNLLVDFDLSCSHWFQEGCLTIFTIGAVIA
VPFMSMLADRYGRKPIIVTTAILAFLANMAASFSPNEAIFLILRAFIGACSDSYLSVA
SVATCEYLSEKRAMITVVNVAMSLGMVTLVLTMTDWRWRVFIISLPGVYGFAL
WYLPESPFWLITTKNTEKELKXIYKTANRMVISLVYFAISFMSVELGGDOVAFLYSS
LIEIPAGLAVIPLMKMGKRMIVMCLVEQTLALIGVTVELDSYEFKLIVMLVAKYMA
TIISVHPIMATEQOPTSVRSLCSLNIPOSGIIMSPLYKHIVMSPNWIPFVYAL
FSFISATLAEFMHETKNKLPDIIESLSYPSSETNDLSAYRRSSSSSVASLSKTSVR
SKTTSSESYSKKLDTKNERTENSRDYKHGTYSTPONCNPFRRHSVLTQRMFI
VNQHPREYTEFLNSIFYQNDCCNNYLMAMTIHHSASERKSIDDSIDPRGEACKOLY
EOLNRNLVSLNSNMOTTFKQWVYETRIGHEISLSSCLOENMSKCTMPYAPHEOLY
QDSAKRNLKQTRLDARQIFDAEDSQSEKLDRAIDSLPLVAIGDNLKIKEL
GAKPKFDAETHTKTDRLAMKGTAEAVKEDTGGLDIDFAETTLNSVNDYFRLYQH
HFDOSLNNLNKNSGIEDEKSLKMGTVFVTKAAEHFRTHIGYSLCKYCDENEAFTRL
DAWKNYIKNGTLEKLEFQOSMENTNNYDLISLKSVTTKAYEEFENTAGPEDEESI
ENIERIIVDGLMESAKKPLSEKRSWVEPPLNRSVAEVMSSLENFAYLLREISHOL
LEIYLEDEGREFMISILKPIEDFYTMRLCRMEDNSTPKVLEYLELRIALAGOLSTIV
EYKDOQTEIKNGLELKNKKMMNERVKDAIRASHREFTDKMPNHSSETYKESADAKS
ALPTPODEVVATQNLHLRLWEELANENTKVALIAHVMIDINESLETPPDDDAIFA
ALDRASHVYKKLLOSINDVWLTDGTATLSKGLVKFEFLCDAEYLRDALVLDLRAGTH
DNLDSITNKLEQLKTMSS"
complement(join(19585..19730,19811..20069,20113..20325,
20573..20758,20806..20952,21015..21091,21143..21350))
/gene="B0361.7"
/map="3"
/note="coded for by C. elegans cDNA CEESW24F; similar to
acid phosphatase"
/codon_start=1
/db_xref="PID:g458955"
/translation="MLLLVLLIGASGINAVYKEVPIQANTDLEVTWVRHGDRT
PAELLFPDITKMEGLGELGEGAQOYRLQWLKRRYGSWGEKFNRAIYIRSSD
YNRITMSQAQNMAGLFPPKYPYIAGLGMOPYPYHTISKPTDKELYEASCPAELEMN
AQWKSTKANGIRKKFARLSFESQKLNLPNMEKATWIRIFDNFCEKONNITWSSWMN
SSIFERVDOIYNEVSOLFETHDTRLRGSTLLEIFHRFSKASGSLKEAKFYAYS
AHDSTIALLATLGVFDIYPKYATCLLIEMHKLANETRLIRYFHKNETDIDRLIYS
IPGDDPCTLOKLGDDLKXKYPEDWEAEVCIISIIIVISTVCSCTMLEVEKQKRLIRF
PVDGLRDTAPMLGGDDSD"
complement(join(21997..22266,22696..22845,22891..23081,
23364..23470,23685..24031,24082..24355,24615..24725,
24773..24849,25101..25395,25508..25636,25682..25777,
25914..26026))
/gene="B0361.8"
/map="3"
/note="similar to cytoplasmic domain of synaptobrevin"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g458956"
/translation="MSDRVLSLISHSITVFEYLVPLIALIIPESLYSGFRRSKTV
AFHPCYCNAGGGERVLMAIRTMQKKFPDHYFVYSGDITDQKQILKARQRIEL
DPSNIQFIYLMWRTLVLEARYKHCTMLFQALAGLILALEAFRMVPAVFIIDSGYPLS
LPARLGSQKVVAVYVHYPTISCDMLDVESROETFNNSSTIASNVLWSGKLTYYRLF
ACLYWLAGRAHVGVMNGSWTQRHITSWSRQDVSIYPPCYEAFNLINESVBSLLE
DTKTVRLISVGOIRPEKNHKLQLEVLHDVKEPBLEKMGYNVELCIAGCGRNEQDERV
MLKNEAEKLDISEQLIWQLVNPEYEDLVLELSKALISIHITMNEHFGISVEAMAAS
ILSNDSGBRMDIVKDYEGHCYGLISTKEEYVETILKIVEGLKRNDRTRYARKSL
TRFGEAEVSRFGKLLKAKNGIFPNRKWKILGTGKSKKCEHFQFOINHITSRFLY
FLYGLFNTMKLYSILVFKHNVDTSDVKLFKSECDLSFSFQSGVQSEFMTAKL
LYERSGLGARSSVKENEYLVHCYVRNDGLSAVCYDAEYQQRVAMSFLGRVLDFTTR
VPATQWPGIRSDKDCSYTGKLDLLEKWNPREADPMTRVQEEVEETKMVMHNTIQSVL
DRGEKLDLVKSENLSDOCKMYTSARKMKCCNTY"
join(28730..28781,29112..29230,29285..29449,29531..29692,
30013..30129,30300..30479)
/gene="B0361.2"
/map="3"
```

```
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g458950"
/translation="MHTFYTHRAIPAIGYGMFRRPDDDDDDYSGRKVIKPKPIAEK
YAKKIGSEFSTGKTFVTGSDSKSQKDFYGSQOVRNDVYKSSDSGAPLTEDENKLSAK
ILKAEMKGDIDLVKLLKRLKLSGISGDDEPPPKSKSEVTMMRRDREGNILPASSRRSD
SDRHGESSRMRREYKESODLDSMVREKGTAGDQLRLFERSLIKSSKIRRHDESV
DDIAEMQKQKKKSDEKDKRKREKESIKGVYKISFEYFLKTFPEI"
complement(join(34166..34284,34390..34719,34823..34958))
/gene="B0361.9"
/map="3"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g458957"
/translation="MEVLSIALISCTTLCATTEWGDLRHLNPARQAPFYDYTDE
KVNVCPOGLHADAIPERYVFGTMTATMTYDEHDOCLQCAEKPRCAVNFHFAYO
KGFCELLGOLDNPSIMRPFKATYERKRCRELDVEDYDEAARPIGSEITEKLPRD
MAREKIDMSKLMKLSAKVKEFNGAGGFPRAAR"
join(36002..36038,36889..36983,38276..38405,39547..39722)
/gene="B0361.1"
/map="3"
/codon_start=1
/evidence=not_experimental
/db_xref="PID:g458949"
/translation="MVTSSPSLHVHYEHKRIERSFDDCSRCIDSSRLKKNHIAVGIN
TYLAIVENDDLDEHLIIVPTQHCSSITQLDENVWDEMRLMRKGLVAVWKSQNRDCIF
```

Note: remainder of annotations omitted.

Query Match 2.7%; Score 23; DB 25; Length 42528;
Best Local Similarity 89.7%; Pred. No. 9.32e-01;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 28059 agaaggaactgaataatgtgtata 28087
||||||| ||||| ||||| |||
QY 199 AGAGGAGAGAACTGAATAATGTGTATA 227

RESULT	10
LOCUS	HEAH10655 819 bp DNA
DEFINITION	Haemophilus influenzae Rd shikimate 5-dehydrogenase (aroE) gene, complete cds.
ACCESSION	L45296 L42023
NTID	91005525
KEYWORDS	aroE gene.
SOURCE	Haemophilus influenzae (strain Rd) DNA.
ORGANISM	Haemophilus influenzae
REFERENCE	Eubacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.
AUTHORS	1 (bases 1 to 819) Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A., Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J.-F., Dougherty,B.A., Merrick,J.M., McKenney,K., Sutton,G., Fitzhugh,W., Fields,C.A., Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.-I., Glodek,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T., Hedblom,E., Cotton,M.D., Utterback,T.R., Hanna,M.C., Nguyen,D.T., Sadek,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L., Fuhmann,J.L., Geoghagen,N.S.M., Gnehm,C.L., McDonald,L.A., Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C.
TITLE	Whole-genome random sequencing and assembly of Haemophilus influenzae Rd
JOURNAL	Science 269 (5223), 496-512 (1995)
MEDLINE	95350630
COMMENT	Address all correspondence to: J. Craig Venter The Institute for Genomic Research 932 Clopper Rd, Gaithersburg, MD, 20878 (New address as of Sept 9, 1995: 9712 Medical Center Dr, Rockville, MD, 20850) Jean-Francois Tomb, Brian A. Dougherty, and Hamilton O. Smith are with the Johns Hopkins University School of Medicine, Baltimore, MD, 21205. Joseph M. Merrick is with the State University of New York, Department of

Microbiology, Buffalo, NY, 14214. Chris Fields' current address is The National Center for Genome Resources, Sante Fe, NM, 87505. Keith Mckenney is with the National Institute for Standards and Technology, Gaithersburg, MD, 20878. All other authors are with The Institute for Genomic Research, 932 Clopper Rd, Gaithersburg, MD, 20878

Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332. e-mail: mark@ember.gatech.edu
Submitted by: Robert D. Fleischmann
e-mail: rdfs@leisctigr.org

Annotation: Owen White

e-mail: owhite@tigr.org

Biological role information, putative identifications, sequence alignments, on-line name and sequence search capability are available at TIGR's World Wide Web site (URL: <http://www.tigr.org/>).

FEATURES

source

CDS

Location/Qualifiers
1..819
/organism="Haemophilus influenzae"
/strain="Rd"
/sequenced_mol="DNA"
1..819
/gene="aroE"
/standard_name="HI0655"
/note="similar to SP:P15770 [Escherichia coli] percent identity: 49.1; corresponds to position 698946 to 698125 of I42023; identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="shikimate 5-dehydrogenase"
/db_xref="PID:g1005526"
/translation="MDLYAVWGNPIAOSKSPLIQNKLAQTHOTMEYIAKLGDLDAFE
OQLAFREEGAKGCNITSPFEKRAYQLADEYSQRAKLAECNTLKKLDDGKLYADNTD
GIGLYDLQRLNMLRPNOHYLIVAGAGATKGVLPLLAQONIVLANRTFSKTELAE
RFQPYGNIQAVSMDSIPLOYDYDLYINATSAAGLSGTAVDAILKLSAFYDMQYAKG
TDTPIFALCKSLGILTNSDGFGLVYAQAASHFHLMRGVMPDFVSVEQLKRAML"

BASE COUNT
ORIGIN

246 a 152 c 184 g 237 t

Query Match

Best Local Similarity 73.9%; Score 22; DB 19; Length 819;

Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Db 570 tggattaagtgggtgaacgcttcagttgatgtgaatttaaa 615

QY 157 TGGATTAATTGGTGAACAGCGCTGTGATGATCCAGAAATTTAGAA 202

RESULT 11

LOCUS MUSG1CYC 912 bp mRNA ROD 11-SEP-1992
DEFINITION Mus musculus D-type G1 cyclin catalytic subunit (PSK-J3/CDK4) mRNA,
complete cds.

ACCESSION
NID L01640
g193363

KEYWORDS D-type cyclin; G1 cyclin; catalytic subunit; protein kinase.

SOURCE Mus musculus (strain BALB/c, sub_species domesticus) (library: WEHI-3 CDNA) CDNA to mRNA.

ORGANISM Mus musculus

Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.

REFERENCE 1 (bases 1 to 912)

AUTHORS Matsushime,H. and Sherr,C.J.

TITLE Identification and properties of an atypical catalytic subunit

(p34psk-j3/CDK4) for mammalian D-type G1 cyclins

JOURNAL Cell (1992) In press

FEATURES location/Qualifiers

source 1..912

/organism="Mus musculus"

/strain="BALB/c"

/sub_species="domesticus"

/cell_type="macrophage-like"

/germline

CDS

/sequenced_mol="CDNA to mRNA"
/tissue_lib="WEHI-3 CDNA"
1..912
/gene="PSK-J3/CDK4"
/codon_start=1
/function="kinase"
/product="D-type G1 cyclin catalytic subunit"
/db_xref="PID:g193364"
/translation="MAATRYEPVAEIGVAGTYKARDPHSGHFAKSVRPNGA
AGGAPVSTVREVALRLRLAEFHPNVRMLDVCATSRDRIKTVLFEEHIDQLRT
YLDKAPPGPLPVETIKDLRQFLSGDLFHLANCIVRDLKPNILVTSNGTVKLADFG
LARIYSQMALTPVVVITWYRAPEVLLQSTYATPVDMVSGCIFAEFRRKPLFCGNS
EADQLGKIFDLIGLPPEDDMPREVSLPRGAFAPRGPRPVQSVPEMESGAQLLEML
TFNDHKRISAFRALQSHLKEESDAE"

BASE COUNT
ORIGIN

206 a 238 c 257 g 211 t

Query Match 2.6%; Score 22; DB 66; Length 912;

Best Local Similarity 86.5%; Pred. No. 4.06e+00;

Matches 32; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 156 cacagttcgtgagtgagc-ctgttaagagagctga 191

QY 712 CACAGTCCAGAGGTGCTCTGCTAAGCAGCGCTGA 748

RESULT 12

LOCUS HIU32694 12500 bp DNA BCT 03-MAY-1996
DEFINITION Haemophilus influenzae aroE, asd, kdtA, kdtB, mdab, rep, tag, yhes,
yhd, yrdC, yrd genes from bases 87095 to 99594 (section 9 of 163)
of the complete genome.

ACCESSION
NID U32694 L42023
g1220729

KEYWORDS complete genome; shotgun sequencing.

SOURCE Haemophilus influenzae.

ORGANISM Haemophilus influenzae

Eubacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.

REFERENCE 1 (bases 1 to 12500)

AUTHORS Tatusov,R.L., Mushegian,A.R., Bork,P., Brown,N.P., Hayes,W.S.,

Borodovsky,M., Rudd,K.E. and Koonin,E.V.

TITLE Metabolism and evolution of Haemophilus influenzae deduced from a

whole genome comparison with Escherichia coli

JOURNAL Curr. Biol. 6, 279-291 (1996)

REFERENCE 2 (bases 1 to 12500)

AUTHORS

Fleischmann,R.D., Adams,M.D., White,O., Clayton,R.A.,

Kirkness,E.F., Kerlavage,A.R., Bult,C.J., Tomb,J.-F.,

Dougherty,B.A., Merrick,J.M., Mckenney,K., Sutton,G., FitzHugh,W.,

Fields,C.A., Gocayne,J.D., Scott,J.D., Shirley,R., Liu,L.-I.,

Glodek,A., Kelley,J.M., Weidman,J.F., Phillips,C.A., Spriggs,T.,

Hedblom,E., Cotton,M.D., Uterback,T.R., Hanna,M.C., Nguyen,D.T.,

Saudex,D.M., Brandon,R.C., Fine,L.D., Fritchman,J.L.,

Fuhrmann,J.L., Geoghegan,N.S.M., Gnehm,C.L., McDonald,L.A.,

Small,K.V., Fraser,C.M., Smith,H.O. and Venter,J.C.

Whole-genome random sequencing and assembly of Haemophilus

influenzae Rd

TITLE Unpublished (1995) see COMMENT for author addresses

REFERENCE 3 (bases 1 to 12500)

AUTHORS White,O.

JOURNAL Direct Submission

REFERENCE Submitted (27-JUL-1995) to GSDB by: Owen White

TITLE Address all correspondence to:

J. Craig Venter

The Institute for Genomic Research

932 Clopper Rd, Gaithersburg, MD, 20878

(New address as of Sept 9, 1995: 9712 Medical Center Dr, Rockville,

MD, 20850)

Jean-Francois Tomb, Brian A. Dougherty, and

Hamilton O. Smith are with

the Johns Hopkins University School of Medicine, Baltimore, MD,

21205.

Joseph M. Merrick is with the State University of New York,

Department

of Microbiology, Buffalo, NY, 14214. Chris Fields' current address is
The National Center for Genome Resources, Sante Fe, NM, 87505.
Keith Mckenney is with the National Institute for Standards and
Technology, Gaithersburg, MD, 20878.
All other authors are with The Institute for Genomic Research, 932
Clopper Rd, Gaithersburg, MD, 20878
Predicted open reading frames were determined using Genemark
software, kindly supplied by Mark Borodovsky,
Georgia Institute of Technology,
Atlanta, GA, 30332. e-mail: mark@amber.gatech.edu
Submitted by: Robert D. Fleischmann
e-mail: rdfleisc@tigr.org
Annotation: Owen White
e-mail: owhite@tigr.org
Biological role information, putative identifications, sequence
alignments, on-line name and sequence search capability are
available at TIGR's World Wide Web site (URL: <http://www.tigr.org/>)
**** Comment added by NCBI staff:
The H. influenzae sequence was re-analyzed with several different
methods for protein similarity detection; the resulting set of
predicted genes that have homologs in databases was used to
re-train the GenMark program in order to predict additional genes
among open reading frames without homologs. As a result, some
predicted small genes from ref. 1 (with subsequent modifications)
whose existence could not be corroborated were discarded, and 23
new genes were predicted. An attempt was made to tentatively
reconstruct all coding regions containing frameshifts or in-frame
stop codons, including those that have been annotated as
misc_features in ref. 1. The sequence was re-annotated based on
this analysis. The DNA replication origin (oriC) was chosen as the
new zero point. Gene names were assigned to those and only those
genes for which apparent orthologs could be identified in E. coli.
Address all correspondence concerning the H. influenzae sequence
re-analysis and re-annotation to:
Roman L. Tatusov or
Eugene V. Koonin
National Center for Biotechnology Information,
National Library of Medicine, National Institutes of Health,
Bethesda, MD 20894
email: tatusov@ncbi.nlm.nih.gov, koonin@ncbi.nlm.nih.gov
R. L. Tatusov, A. R. Mushegian, K. E. Rudd, and E. V. Koonin are with the
National Center for Biotechnology Information,
National Library of Medicine,
National Institutes of Health,
Bethesda, MD 20894
W. S. Hayes and M. Borodovsky are with the
School of Biology, Georgia Institute of Technology, Atlanta, GA
30332-0230
P. Bork and N. P. Brown are with the
European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900,
Heidelberg, Germany.

NCBI gi: 1220729
Location/Qualifiers
1..12500
/organism="Haemophilus influenzae"
/strain="Rd"
complement(54..1169)
/gene="asd"
/standard_name="HIN0066"
/note="homolog of DHAS_ECOLI, BLAST score: 704; HI0646;
NCBI gi: 1220730"
/codon_start=1
/transl_table=11
/product="aspartate-semialdehyde dehydrogenase"
/db_xref="PID:g1220730"
/translation="MKNVGFIGRWGVSVLMDRMSQENDFENLPVFTTSQAGQKA
PVEGKDGADLKSADFIEELKLDIYTCOGDYTNEYYPKLKATGWDGYWDAASAL
RMKDAIIVLDPVNOHVISEGLKGTFTVGGNCTVSLMLAIGLEKDLVEMISVA
TYQASGAGAKNMRRLSQMGLQAVSELKDPASSILDIERKVTAKMRADNFPITD
FGALGGSLIPWIDKILPETGOTKEEKGYAETNKILGSDNPPIPVDCIVRIGALRC
HSQAFITIKLKKDLPLEIEIQIASHNEMVKVIPNDKEITLRELTPAKVTGTLSPVGR

CDS
LRLKAMPPEYLAFTVGDQLLWGAAEPVRRILKQLVA"
complement(1328..2044)
/gene="yhd"
/standard_name="HIN0067"
/note="homolog of SSU20224_2, BLAST score: 99; HI0647;
NCBI gi: 1220731"
/codon_start=1
/transl_table=11
/product="Mg2+-transporting ATPase"
/db_xref="PID:g1220731"
/translation="MENSLLLTALFNPDLIIFSKMLLAVLGSVIGLERLKKPV
GVKCAIIAVTTCVLTIVSIOAEHYAQVSENIRTPMRLLAAQVIGIGFLGAVILH
KKNDASGLTTAAITWASAGIGIAGAGFVDAVIAVMILVSIIRLSPVQRVHRKS
QRRRTFNILVNDASIGKVTQLLVNQYRIEHIQVQDSSGSEVRLQRCFSIDSTML
KDAYALLKAEADVISEVDN"
complement(2322..2948)
/gene="mdaB"
/standard_name="HIN0068"
/note="homolog of MDAB_ECOLI, BLAST score: 633; HI0648;
NCBI gi: 1220732"
/codon_start=1
/transl_table=11
/product="modulator of drug activity"
/db_xref="PID:g1220732"
/translation="MATPEFFIICFRSKMNLIDGKAFGSHGELNHTLKKAK
EYTLAIGHNVKEIVDAGYDVEAEIEKFLMDAVIWMQPSWMHHEPTVKYIDEVLT
NGHCKLYHSDGRHSVNPTEGYGTGILLQCKHMLSTWNPATIAFTRGDEFEKGVD
VLYMHFHLNEFLGLRLPTFLCNDVVKSPQVEGYLADYQAHLEKVEG"
complement(3190..5202)
/gene="rep"
/standard_name="HIN0069"
/note="homolog of REP_ECOLI, BLAST score: 1657; HI0649;
NCBI gi: 1220733"
/codon_start=1
/transl_table=11
/product="DNA helicase"
/db_xref="PID:g1220733"
/translation="MKLNPOOQGAVEYVTPCLVLGAGSGKTRVIINKIAHLIEKCG
YSPKQIAAVTFTNKAAREMKERYAHSIGKEQSGLLVSTFHTLGDILKREYKALGFK
SNMTLFDEHDOFALLKELTADYLEDKDLRELISVSNMKNLISPQAFALARAK
YOTFAKCYERYAQIRTYNALDDDDLIMPTLFEKONEEVRSKWAKIRYLLVDEYQD
TNTSQYELIKLLVGDRACTVVGDDDSIYSWRGARPEMVRRLRDPFLNVIKLEON
YRSTORILHCANLIDNNEHVEFKLFSTIGKEKLLVIEAKNEHEAERIVAELIAH
RESRKTXYKDIAILIRGNHQSRLLEKVLQMNRIPYKISGTSFESRAEIKDMAYLRL
VYNQDDDAFLRIYVNTPKREIGTATLQKLGLOEKHISLFEATFEFELLQRTPKAY
DSLQKGRWIVEINDEIQRSEPRAVRSMLSAIEEYLYEYATSPKAEMQSKNVA
LEDWVADMLKDETNEMNLNQVTRTLRLDMLERGEDDDSDQVDFLTHASKGLEF
PYVYLIGMEGILPHQTSIDEDNVEERRLAYGITRAQKELTSLCRERROYGELVR
PEPSRFLAELPNDVDVLMERDKPLTTEQKQEKTONQDLRLRAILKS"
complement(5208..5420)
/standard_name="HIN0070"
/note="weak similarity to YEJM_ECOLI, BLAST score: 60;
HI0650; NCBI gi: 1220734"
/codon_start=1
/transl_table=11
/product="H. influenzae predicted coding region HIN0070"
/db_xref="PID:g1220734"
/translation="MIKIFITLALIVLSGCGSVKLLIDPTEKYTAYAGVADLEMAQ
QWGLPIIDLPISFLDVLTPYAWAQ"
complement(5417..5887)
/gene="kdtB"
/standard_name="HIN0071"
/note="homolog of KDTB_ECOLI, BLAST score: 441; HI0651;
NCBI gi: 1220735"
/codon_start=1
/transl_table=11
/product="lipopolysaccharide core biosynthesis protein
(cytidilyltransferase ?)"
/db_xref="PID:g1220735"
/translation="MTSVIYEGTFDPITNGHLDIERSAVIPEPRVLAVAVNSPKPL
FSLEREVELYRQSVHLSNVEYGFSDLLANVIKQHNISALIRGVRTTDFEYELQLA
ALNRLTKGVDSLFFPPAEKMAFVSSITVREIYVLAGDVAELVPPVFNALKAR"
complement(5884..7167)

CDS

FT /db_xref="PID:g1220731"
FT /translation="MENSLLLTALFNPDHIIIFSCKMLAMVLGSGVIGLERELKRPV
G VKTCALIAVTTCVLTIVSIQAAEHYAQVSENIRTDPMRLAAOYISGIGFLGAVILHK
K NDAISGLTTAIIWASAGIGIAGAGFVFDVAVIATVMIIVSIRLSPVQRMVHRKSQR
R RTKFNIILVNDAESIGKYTQLLVNNQYRIEHIOVKDQSSGEVRLQIRCFSIDSTMLKDA
Y ALKAEEDGVISVEVDN"
FT complement(2322..2948)
FT /gene="mdab"
FT /standard_name="HIN0068"
FT /note="homolog of MDAB_ECOLI, BLAST score: 633; HI0648;
FT NCBI g1: 1220732"
FT /codon_start=1
FT /transl_table=11
FT /product="modulator of drug activity"
FT /db_xref="PID:g1220732"
FT /translation="MATPEFFIILICFRSKKMILLDDGKAFGSHGELNHTLHKKAK
E VLTALGHNVKETVIDAGYDVEAEIEKFLMMDAVIMQMPSTMHEPWTVKKYIDEVLTN
G HGKLYHSDGRHSVNPTREGYGTGLLGKHKMLSLTWNAPIEAFTRGDFEFGKGVDL
Y MHEKLNELGLTRLPTFLCNDVKSPOVEQYLADYQAHLEKVEG"
FT complement(3190..5202)
FT /gene="rep"
FT /standard_name="HIN0069"
FT /note="homolog of REP_ECOLI, BLAST score: 1657; HI0649;
FT NCBI g1: 1220733"
FT /codon_start=1
FT /transl_table=11
FT /product="DNA helicase"
FT /db_xref="PID:g1220733"
FT /translation="MKLNPOQQQAAVEYVTGPCLVLAGAGSGKTRVLIINKIAHLIEKCG
Y SPKOIAAVTFTNKAREMKERVASHSIGKEQSKGLLVSTFHTLGFDLKREYKALGFKS
N MTFEDENHDFALLKELTADVLKEDKDLRELISVISNKNNDLISPKQAFALARDAKYO
T FAKCYERYATQIRTYNALDFDDLIMLPTLLFKONEEVRASKWQAKIRYLLVDEYQDINT
S QVELIKLLVGDRACTVVGDDDSISWRGARPENNVRLRDDFRLNVIKLEQNYRST
Q RILHCANILIDNNEHVEDKLLFTIGKEKLLVIEAKNEHEAERIYAEIIAHRFSRK
T KYKDYAILYRGNHOSRLLKVLQMNRIPYKISGTSFSSRAEIKDMMAVYLRVNNQDD
D AAFRLRIVNTPKREIGTATLQKLGEIAOEKHISLFEAIEFELIQIRITPKAYDSLQKFG
R WIVELNDEIQRSEPERAVRSMLSAIIHEEYLYEYATSPKAEMQSKNVATLFDWVADM
L KGEETNEPMNLNOVYTRLTLRDMLERGEDDDSDQVQMLTLHASKGLEFPYVYLIGME
E GLIPHOTSIDEDNVEEERRLAYVGITRAQKELTFSLCRERROYGELVAPPEPSRFLAEL
P NDVLMERDKPKLTTEQKOEKQTONOLDRLRAILKS"
FT complement(5208..5420)
FT /standard_name="HIN0070"
FT /note="weak similarity to YEJM_ECOLI, BLAST score: 60;
FT HI0650; NCBI g1: 1220734"
FT /codon_start=1
FT /transl_table=11
FT /product="H. influenzae predicted coding region HIN0070"
FT /db_xref="PID:g1220734"
FT /translation="MIKIFIFLALIVLSGGSVYKLIDPTEKYTAYAGVADLEMAQ
O WGLPILDPLSFLDITVLIPYAWAQ"
FT complement(5417..5887)
FT /gene="kdtb"
FT /standard_name="HIN0071"
FT /note="homolog of KDTB_ECOLI, BLAST score: 441; HI0651;
FT CDS

FT NCBI g1: 1220735"
FT /codon_start=1
...
Note: remainder of annotations omitted.
Query Match 2.68; Score 22; DB 10; Length 12500;
Best Local Similarity 73.9%; Pred. No. 4.06e+00;
Matches 34; Conservative 0; Mismatches 12; Indels 0; Gaps 0;
Db 8749 tttaaatttcagcatcacactgaagccgttcacaccatcaatcca 8794
CP 202 TTCTAAATTTCTCGATCATCCAGCCCTGTTCACCATATATTC 157
RESULT 14
LOCUS HIU32694 12500 bp DNA BCF 29-FEB-1996
DEFINITION Haemophilus influenzae aroe, aed, kdtA, kdtB, mdab, rep, tag, yhes,
yhd, yrdC, yrd genes from bases 87095 to 99594 (section 9 of 163)
of the complete genome.
ACCESSION U32694 L42023
MID 91220729
KEYWORDS complete genome; shotgun sequencing.
SOURCE Haemophilus influenzae.
ORGANISM Haemophilus influenzae
Eubacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
Haemophilus.
REFERENCE 1 (bases 1 to 12500)
Fleischmann, R.D., Adams, M.D., White, O., Clayton, R.A.,
Kirkness, E.F., Kerlavage, A.R., Bult, C.J., Tomb, J.-F.,
Dougherty, B.A., Merrick, J.M., McKenney, K., Sutton, G., FitzHugh, W.,
Fields, C.A., Gocayne, J.D., Scott, J.D., Shirley, R., Liu, L.-I.,
Glodek, A., Kelley, J.M., Weidman, J.F., Phillips, C.A., Spriggs, T.,
Hedblom, E., Cotton, M.D., Utterback, T.R., Hanna, M.C., Nguyen, D.T.,
Saudek, D.M., Brandon, R.C., Fine, L.D., Fritchman, J.L.,
Fuhrmann, J.L., Geoghegan, N.S.M., Gnehm, C.L., McDonald, L.A.,
Small, K.V., Fraser, C.M., Smith, H.O. and Venter, J.C.
Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd
Unpublished (1995) see COMMENT for author addresses
2 (bases 1 to 12500)
White, O.
Direct Submission
Submitted (27-JUL-1995) to GSDB by: Owen White
3 (bases 1 to 12500)
Tatusov, R. L., Mushegian, A. R., Bork, P., Brown, N. P., Hayes, W. S.,
Borodovsky, M., Rudd, K. E. and Koornin, E. V.
Metabolism and evolution of Haemophilus influenzae deduced from a
whole genome comparison with Escherichia coli
Curr Biol., in press (1996) In press
Address all correspondence to:
J. Craig Venter
The Institute for Genomic Research
932 Clopper Rd, Gaithersburg, MD, 20878
(New address as of Sept 9, 1995: 9712 Medical Center Dr, Rockville,
MD, 20850)
Jean-Francois Tomb, Brian A. Dougherty, and
Hamilton O. Smith are with
the Johns Hopkins University School of Medicine, Baltimore, MD,
21205.
Joseph M. Merrick is with the State University of New York,
Department
of Microbiology, Buffalo, NY, 14214. Chris Fields' current address
is
The National Center for Genome Resources, Sante Fe, NM, 87505.
Keith Mckenney is with the National Institute for Standards and
Technology, Gaithersburg, MD, 20878.
All other authors are with The Institute for Genomic Research, 932
Clopper Rd, Gaithersburg, MD, 20878
predicted open reading frames were determined using Genemark
software, kindly supplied by Mark Borodovsky,
Georgia Institute of Technology,
Atlanta, GA, 30332. e-mail: mark@amber.gatech.edu
Submitted by: Robert D. Fleischmann

e-mail: rdfs@ctigr.org
Annotation: Owen White
e-mail: owite@ctigr.org
Biological role information, putative identifications, sequence alignments, on-line name and sequence search capability are available at TIGR's World Wide Web site (URL: <http://www.tigr.org/>)
**** Comment added by NCBI staff:

The H. Influenzae sequence was re-analyzed with several different methods for protein similarity detection; the resulting set of predicted genes that have homologs in databases was used to re-train the GenMark program in order to predict additional genes among open reading frames without homologs. As a result, some predicted small genes from ref. 1 (with subsequent modifications) whose existence could not be corroborated were discarded, and 23 new genes were predicted. An attempt was made to tentatively reconstruct all coding regions containing frameshifts or in-frame stop codons, including those that have been annotated as misc_features in ref. 1. The sequence was re-annotated based on this analysis. The DNA replication origin (oriC) was chosen as the new zero point. Gene names were assigned to those and only those genes for which apparent orthologs could be identified in E. coli. Address all correspondence concerning the H. Influenzae sequence re-analysis and re-annotation to:

Roman L. Tatusov or
Eugene V. Koonin
National Center for Biotechnology Information,
National Library of Medicine, National Institutes of Health,
Bethesda, MD 20894
email tatusov@cbl.nlm.nih.gov, koonin@cbl.nlm.nih.gov R. L. Tatusov, A. R. Mushegian, K. E. Rudd, and E. V. Koonin are with the National Center for Biotechnology Information,
National Library of Medicine,
National Institutes of Health,
Bethesda, MD 20894
W. S. Hayes and M. Borodovsky are with the
School of Biology, Georgia Institute of Technology, Atlanta, GA 30332-0230
P. Bork and N. P. Brown are with the
European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900, Heidelberg, Germany.

FEATURES

source

CDS

/organism="Haemophilus influenzae"
/strain="Rd"
/complement(54..1169)
/gene="asd"
/standard_name="HIN0066"
/note="homolog of DHAS_ECOLI, BLAST score: 704; HI0646"
/codon_start=1
/transl_table=11
/product="aspartate-semialdehyde dehydrogenase"
/db_xref="PID:g1220730"
/translation="MKNVGFIGRWGVSVLMDRMSGENDFENLNPVEFTTSQAGOKA
PVFGKDGADLKSADFIEELKLDIIVTCGGDYTNEYVPLKATGWDGYVDAASAL
RMKDAIIVLDPVNQHVISEGLKGIKTFVGGNGCTVSLMLAIGLFEKDLVEMISA
TYQASGAGAKNMRRLSQMGLEQAVVSELKDPASSILDIERKVTAKMRADNFPD
FGAALGSLIPWIDKLPETGOTKEEMKGYAETNKILGSDNPPIVDGLCVRIGALRC
HSQAFYIKLKKDLPLEIEIOIASHNEWVKVIPNDKETLRELTPAKVTGTVSPVGR
LRKLMGPYLAFTVGDQLMGAAEPVRRILKQLYA"
complement(1328..2044)
/gene="yhd"
/standard_name="HIN0067"
/note="homolog of SSU20224_2, BLAST score: 99; HI0647"
/codon_start=1
/transl_table=11
/product="Mg2+-transporting ATPase"
/db_xref="PID:g1220731"
/translation="MENSLLTALFNPDLIIFSKMLAVLGVIGLERELKRPV
GVKCAIAVTTCVLTIVSIQAEHYAOVSENIRTDPMRLAQAIVSGIGFAGVILH
KKNDASGLITAAIIVASAGIGIAGAAGFVDAVIAIVILVSIKPLVDRVHRKS
QRRRTKFNILVNDASIGKVTQQLVNNQYRIEHIQKQSSGEVRLQIRCFSIDSTML
KDAYALKAEDGVISVEVDN"
complement(2322..2948)

CDS

CDS

CDS

CDS

/gene="mdab"
/standard_name="HIN0068"
/note="homolog of MDAB_ECOLI, BLAST score: 633; HI0648"
/codon_start=1
/transl_table=11
/product="modulator of drug activity"
/db_xref="PID:g1220732"
/translation="MATPFEFIIICFRSKNNILLDDGKAFGSHGELNHTLHKRAK
EVLTAIGHNDKVTIDAGYDVEAEIEKFLMMDAVIWMQPSWMMHEPTVKYIDEVLT
NGHKILYHSDGRHSVNPTEGXTGGLGKHKMLSTWNPATEATREGDFEGKGV
VLYMFEHLNFEFLGTRLPFTFLCNDVKSPOVEQYLADYQAHLEKVEG"
complement(3190..5202)
/gene="rep"
/standard_name="HIN0069"
/note="homolog of REP_ECOLI, BLAST score: 1657; HI0649"
/codon_start=1
/transl_table=11
/product="DNA helicase"
/db_xref="PID:g1220733"
/translation="MKLNPOQOQAVEYVTGCLVLAGAGSGKTRVIINKIAHLEKCG
YSPKQIAAVFTNKAREMKERVAHSIGKEGSKLLVSTFHTLGFDLIREYKALGFK
SNMTLFDEHDOFALKRELTADVLEKDLRELISVSNMKNLDSKQAFALARDAX
YOTFAKCYERYATQIRYNALDFDDLIMPLTLFKONEVRSKQAKIRYLLVDEYOD
TNTSQYELIKLVGDRACFTVWGDDOSIYWRGARNVRLRDDPRLNVIKLEON
YRSTQRIHLCANILIDNNEHVDFKKLFSTIGKEKLVIEAKNEHEAERIVAEILAH
RFSRRTKYKDVAILYRGNHQSRLLEKVLQMNPYKISGTSFSFRAEIKDMAYLR
LVNQDDDAFLRIVNTPKREIGTATLQKGLAEQKHSLSFEAIFEFELIQRITPKAY
DSLQKFGRWIVELNDEIQSEPERAVRSMLSAIIHYEYLDEYATSPKAAEMQSNVAT
LEFDVADMLKGETNEPMNLNOVTRLTLRDMLEGEEDDDQVQVMTLHASKGLEF
PYVYLLIGMEEGILPHQTSIDEDNVEEBRLAYVGITRAQKELTFSLCRERROYGELVR
PEPSRFLAELPNDVLEMRDKPKLTTEQKEKTONQLDRLRAILKS"
complement(5208..5420)
/standard_name="HIN0070"
/note="weak similarity to YEJM_ECOLI, BLAST score: 60;
HI0650"
/codon_start=1
/transl_table=11
/product="H. Influenzae predicted coding region HIN0070"
/db_xref="PID:g1220734"
/translation="MKRIFIFLALIVLSCGGSVVKLIDPTERYAYAGVAYDLEMAQ
QWGLPILDPLSFLDVLVLLPYAWAQ"
complement(5417..5887)
/gene="kdtb"
/standard_name="HIN0071"
/note="homolog of KDTB_ECOLI, BLAST score: 441; HI0651"
/codon_start=1
/transl_table=11
/product="1ipopolysaccharide core biosynthesis protein
(cytidyl)ltransferase ?)"
/db_xref="PID:g1220735"
/translation="MTSVYXPGTFDPITNGHLDIERSAVIFPRVLAVANSRSPKPL
FSLSEVELVRQSVHLSNVEVFGFSDLANVIKQHNISAIIRGVRTTDFEYELQLA
ALNRLLTKGVDSLFFPRAEKWAFVSVTIVREITLHGDDVAELVPVPFNALKAR"
complement(5884..7167)
/gene="kdtA"
/standard_name="HIN0072"
/note="homolog of KDTA_ECOLI, BLAST score: 914; HI0652"
/codon_start=1
/transl_table=11
/product="3-deoxy-d-manno-oculosonic-acid transferase"
/db_xref="PID:g1220736"
/translation="MWRFEYTSLLILICQPLICFIGLLSVKSPRYRORLAERYGYGN
ASCPPOGIFIHAASVGEVIAATPLVROLQODPHLSITFTTPTPSERVKATFGDS
VFHYLLPLDLPFSIHRTNFVQPKLCIVMETELPNLIHQFLRNIPVIANARLSAR
SAHRGKIKAHLOTWMSQISLIAAODNISGRYATLGYPKEKLNITGNKYDLNDE
LKRKIDSLRLWKQDRPIWIAASTHNGEDEIILKSHRALLAKYPNLLLVPRHPERF
NNVADLLKEKFOFIRSTNELNENTOVILGDSMGEIMMYGISDIAFVGSLVKHG
GHNPLEPLAFKMPVITGHTFNPPEIFRMLVEYQGVLEVNSTADALERAVEALLNSKE
SRERLGNAGYEVLNENRQALQRLDLLKPYLERNY"

Note: remainder of annotations omitted.

Query Match

2.6%; Score 22; DB 20; Length 12500;

CDS

Search completed: Mon Feb 10 23:54:38 1997
Job time : 637 secs.

CDS

```
/codon_start=1
/product="R06F6.5"
/db_xref="PID:g577765"
/translation="MPSHLNQNTSGRHNMDLNSSISNFGTPEVQSTPALLEGKRA
TVPSSYASPLNTASAPCSDFAVSAPAVPQHLKDTPGCSVHMSPSLVQSGEKSAQ
TQNTPANLSFGNSSSFAPTKPAPOSIQTSFGQAMHAPLRLDRKVEPAKISR
NTFTARSTPLSTPTQRTVTSRLAEAEQPMEEADADTWTVTFGQPSQVSILNLF
SRHGEVSSHOTPSKGNFIHMRYSCTVHAQOASRNGLDQDTFIGVQCTNKDING
SASGIVARSSNIAAANRASAMYSNFVENDMADQSVNHNENSVLNSNVFDANNSLNS
SRISVRSGVMRPLADQRTNGTSPSRKAPDGLNKFMTIGLN"
join(20156..20257,20501..20589,22181..22430,22479..22615,
22668..22989,23996..24125,24266..24351,24417..24551)
/note="homeobox protein, which possibly is the homolog of
the drosophila rough gene (Swiss Prot accession number
P10181)"
```

CDS

```
/codon_start=1
/product="R06F6.6"
/db_xref="PID:g577766"
/translation="MHRKGGQTLRLYTHFLHLSPHHSLDSIMORKMVPSSAFHP
YTRPIPTSGHDLPVDSLRLLLSAEALIALQADASKILPSYEPVKDFSPQLPAM
TPTPIATPSIQEQPQLOSPSPAPNEKSRKRRTFSPQATRLAEYIGDSYAREKR
HLAQSLKLSNQVKTWQNRRAKDRDRKSENASNHTSNRRSSPSRKSSTPTP
TQATQFDMPTQIQTASPTTADSAIFPTSPESIIOKIEQFPNQILPNFDILQTYLQ
SLSSQIPLOQFVSTPPLPDNMLYHVFIFESQETCSVSFORNSAMRSLIFV
FFATIFAVDAQYDSADSSQSEDLPDIAPEVSSEIEIVDVNIEGSGDGPDIIDQVH
MFCILPGPGDIRKKRGILEILGRK"
join(25838..25889,26496..26611,26664..26858)
/codon_start=1
/product="R06F6.7"
/db_xref="PID:g577767"
/translation="MRSWIPLLVLFVAVLAVFAQAGKSESEDSRRRPSKSESSESDSDS
KSSDSSSSSESSGDVPSAPNTDSTPVEILAAKPDGILGPDNRVKKRDGLPSF
YDIRKKRGLPSAYDIRK"
complement(join(28010..28079,28137..29078,29714..29846,
29895..30012,30062..30281,30682..31664,32263..32554,
32651..32943,33071..33608,33689..33780,33835..34296,
34347..34457,34507..35132..35346,37261..37597,
37771..37877,37930..38018,38064..38506,38552..38729,
38780..38841))
/note="CDNA EST CEEST88F comes from this gene; CDNA EST
yk109.3 comes from this gene; CDNA EST yk21h1.3 comes
from this gene; CDNA EST yk109.5 comes from this gene;
CDNA EST yk21h1.5 comes from this gene; CDNA EST yk47c1.5
comes from this gene; CDNA EST yk47c1.3 comes from this
gene; CDNA EST yk88h2.5 comes from this gene; CDNA EST
yk92d6.5 comes from this gene; CDNA EST yk92d6.3 comes
from this gene"
```

CDS

```
/note="CDNA EST CEEST88F comes from this gene; CDNA EST
yk109.3 comes from this gene; CDNA EST yk21h1.3 comes
from this gene; CDNA EST yk109.5 comes from this gene;
CDNA EST yk21h1.5 comes from this gene; CDNA EST yk47c1.5
comes from this gene; CDNA EST yk47c1.3 comes from this
gene; CDNA EST yk88h2.5 comes from this gene; CDNA EST
yk92d6.5 comes from this gene; CDNA EST yk92d6.3 comes
from this gene"
/codon_start=1
/product="R06F6.8"
/db_xref="PID:g577768"
/translation="MSEKENPASPTTGREEEIPGSSPEGVYPADEDHIFHEDQAPLR
VESAKHEEIVEEQOQPPEDLEQDMIVEDQOQFMNVQITQEDMYEAGFDVEAGFE
LNHLTEQLNIVVAISQOQAKONEQEHNEEVYEDGSHHMHVHEMIENQFEGDNG
PEEYDNGQOIIDNAMHILITNDGVNITSKQOQFVYVSPSEIANLNLIDNLSTENVH
QVYQALPPIKEKAQDSAYNDQASTSYHHHHEQLDAGKSTRPIIGETVQIRTADG
RLQDAVVKYVRGDEYKIQIMNGEFAYATIDQMLVYQDRSDHEYQOYAPVLLRTD
MASVRNAAQKRSANDLCPVYLKSYQLAPVVDGPHLVHTPNFCCPICDKKYYQKEPS
YIVIRLPACDSCTHKGRSRPVSVCNKYNGTIGILVPVKILKTGOQNSRSRENTPK
TGQRDKRDTGIDTLVHGNYKMHTNSTNDRPSYQLPKHEKSTADIKSIVANRRR
LIAVATNDAIYIWLNPQLLCSVVIDANFKERGELKEIYWKPDSTSIAVTTNQCK
ILYINLDRDDEQCYNETDSADPYFORNSPELFKGSRTAHLHPITILINLADIPQCC
VPSRDEFIVLCQNGFTHTHTWTEGELIASLSFRASSIPFSVDQLOKSENITSKSTYIF
DAVYAPLLGGFAIVLSDQGALLTSNDPNFAPNALIGWAPNMKDATCCDVNHKFLLI
```

Note: remainder of annotations omitted.

```
Query Match      2.6%: Score 22; DB 26; Length 41452;
Best Local Similarity 92.3%: Pred. No. 4.06e+00;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 11613 aataaacagaaaatggcgatatatt 11638
|||||
CP 960 AATAAACAGAAACTGGCCATATT 935
```



[W] [E] [R] [H] (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_n n.a. - n.a. database search, using Smith-Waterman algorithm
Run on: Mon Feb 10 23:54:57 1997; MasPar time 88.81 seconds
833.805 Million cell updates/sec
Tabular output not generated.

Title: >US-08-674-311-1
Description: (122-970) from US08674311.seq (2 of 2)
Perfect Score: 849
N.A. Sequence: 122 ATGGCTCTGGCACCACCAC.....CTGTTTATTACCAAGACAT 970
Comp: TACCGAGAGACCGTGTGTGTC.....GACAAATAATGTTCTGTGA

Scoring table: TABLE default
Gap 6

Nmatch STD : Dbase 0; Query 0

Searched: 113505 seqs, 43611913 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: n-geneseq25
1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7
8:part8 9:part9 10:part10 11:part11 12:part12 13:part13
14:part14 15:part15 16:part16 17:part17 18:part18
19:part19 20:part20 21:part21 22:part22

Statistics: Mean 8.967; Variance 4.877; scale 1.839

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description	Pred. No.	
1	240	28.3	2763	19	T15167	Methylthioadenosine-p	9.50e-159
2	240	28.3	2763	15	Q92813	Human MTase.	9.50e-159
3	240	28.3	2784	16	Q99202	Pseudomonas putida me	9.50e-159
4	73	8.6	1047	2	Q10572	Human Natriuretic Pep	2.34e-33
5	64	7.5	1047	2	Q10572	Human Natriuretic Pep	4.78e-27
6	44	5.2	91	9	Q51746	Oligonucleotide probe	1.45e-13
7	42	4.9	91	9	Q51746	Oligonucleotide probe	2.82e-12
8	40	4.7	204	1	N81164	Base substituted E.co	5.29e-11
9	37	4.4	204	1	N81164	Base substituted E.co	4.02e-09
10	31	3.7	114	12	Q70469	Generic DNA sequence	1.70e-05
11	30	3.5	114	12	Q70469	Generic DNA sequence	6.50e-05
12	30	3.5	114	12	Q70465	Generic DNA sequence	6.50e-05
13	30	3.5	114	12	Q70468	Generic DNA sequence	6.50e-05
14	30	3.5	114	12	Q70467	Generic DNA sequence	6.50e-05
15	30	3.5	114	12	Q70470	Generic DNA sequence	6.50e-05
16	30	3.5	114	12	Q70467	Generic DNA sequence	6.50e-05
17	30	3.5	114	12	Q70465	Generic DNA sequence	6.50e-05

18	28	3.3	114	12	Q70466	Generic DNA sequence	9.06e-04
19	28	3.3	114	12	Q70468	Generic DNA sequence	9.06e-04
20	28	3.3	114	12	Q70466	Generic DNA sequence	9.06e-04
21	28	3.3	114	12	Q70470	Generic DNA sequence	9.06e-04
22	27	3.2	66	21	T13585	TSAR-9 library genera	3.29e-03
23	27	3.2	69	21	T13583	TSAR-9 library genera	3.29e-03
24	25	2.9	74	21	T13613	DC43 TSAR library gen	4.09e-02
25	25	2.9	75	21	T13612	DC43 TSAR library gen	4.09e-02
26	25	2.9	81	21	T13611	DC43 TSAR library gen	4.09e-02
27	25	2.9	82	21	T13610	DC43 TSAR library gen	4.09e-02
28	25	2.9	114	12	Q70473	Generic DNA sequence	4.09e-02
29	25	2.9	114	12	Q70472	Generic DNA sequence	4.09e-02
30	24	2.8	82	21	T13610	DC43 TSAR library gen	1.39e-01
31	24	2.8	114	12	Q70472	Generic DNA sequence	1.39e-01
32	24	2.8	114	12	Q70471	Generic DNA sequence	1.39e-01
33	24	2.8	114	12	Q70471	Generic DNA sequence	1.39e-01
34	24	2.8	114	12	Q70473	Generic DNA sequence	1.39e-01
35	23	2.7	74	21	T13613	DC43 TSAR library gen	4.63e-01
36	23	2.7	75	21	T13612	DC43 TSAR library gen	4.63e-01
37	23	2.7	81	21	T13611	DC43 TSAR library gen	4.63e-01
38	23	2.7	565	6	Q35072	HCV envelope region n	4.63e-01
39	22	2.6	65	21	T13586	TSAR-9 library genera	1.50e+00
40	22	2.6	68	21	T13584	TSAR-9 library genera	1.50e+00
41	22	2.6	69	21	T13583	TSAR-9 library genera	1.50e+00
42	22	2.6	498	3	N50034	Sequence encoding new	1.50e+00
43	22	2.6	501	3	N50030	Sequence encoding new	1.50e+00
44	22	2.6	565	6	Q35072	HCV envelope region n	1.50e+00
45	22	2.6	657	7	Q43519	Degenerate FMN reduct	1.50e+00

ALIGNMENTS

RESULT 1
ID T15167 standard; DNA; 2763 BP.
AC T15167;
DE 29-JUN-1996 (first entry)
KW Methylthioadenosine-phosphorylase gene.
KW Human; methylthioadenosine-phosphorylase; chromosome walking;
KW Interferon-alpha; CDK4I; tumour suppressor; chromosome-9p21;
KW cyclin-dependent kinase-4-inhibitor; probe; primer; mutation;
KW leukaemia; melanoma; diagnostic; dysplastic nevus syndrome; glioma;
KW non-small cell lung carcinoma; cancer; gene therapy; antisense;
KW ribozyme; antibody; imaging; ss.
OS Homo sapiens.
FH Key
FT exon Location/Qualifiers
FT /tag- b 422..615
FT intron /tag- c 616..720
FT exon /tag- d 721..963
FT intron /tag- e 964..1203
FT exon /tag- f
FT /tag- f
PN WO9528169-A1.
PR 12-APR-1995; U04655.
PR 14-APR-1994; US-227800.
PA (REGC) UNIV CALIFORNIA.
PI Carson DA, Nobori T;
DR WPI: 95-373630/48.
PT Cyclin dependent kinase inhibitor gene, related vectors and
PT antibodies - useful for diagnosis, assessing predisposition and
PT treatment of cancers
PS Example 1; Page 96-101; 129p; English.
CC The sequence encodes a methylthioadenosine-phosphorylase, and is
CC located at chromosome-9p21. A cyclin-dependent
CC protein-kinase-4-inhibitor (CDK4I) tumour suppressor gene
CC (T15157-58) is located between this gene and an interferon-alpha
CC gene cluster, and has been isolated by chromosome walking. The
CC CDK4I gene, probe and primer derivatives and the gene product may
CC be used in diagnosis of cancer, particularly melanoma (especially

CC dysplastic nevus syndrome), glioma, non-small cell lung carcinoma
CC or leukemia. The gene may also be used in cancer gene therapy, or
CC in antitumor antisense oligonucleotide or ribozyme construction.
CC Antibodies against CDK4I may be used in diagnosis or in vivo imaging.
SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T;

Query Match 28.3%; Score 240; DB 19; Length 2763;
Best Local Similarity 99.2%; Pred. No. 9.50e-159; Indels 0; Gaps 0;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggttctatagaagactgctaagaagctaggactccggtgccaactcaaaaggagacaatgg 1021
|||||
QY 570 AGGTTCTTATAGAGACTGCTAAGAGCTAGGACTCCGGTGCCACTCAAAAGGGACAATGG 629
tcacaatcgaggagacctgcttagctcccgaggagaaagcttcattccgacactggg 1081
|||||
QY 1022 tcacaatcgaggagacctgcttagctcccgaggagaaagcttcattccgacactggg 1081
|||||
QY 630 TCACAATCGAGGAGGAGCTGTTTACCTCCGGCAGAAAGCTTCATGTTCCGACCTGGG 689
1082 gggcgagatgtatcaacatgaccacagttccagaggtgtctgtctaaggagctggaa 1141
|||||
QY 690 GGGCGGATGTTATCAACATGACCACAGTTCAGAGGTGTTCTTGAAGGAGGCTGGA 749
1142 ttgttaccgaagatcgccatgggacagattatgactgtggaagagcagaggaag 1201
|||||
QY 750 TTTGTTACGCAAGTATCGCCATGCGCAGACAGATTATGACTGCTGGAAGAGCAGAGGAG 809
1202 cagt 1205
|||||
QY 810 CAGT 813

RESULT 2
ID Q92813 standard; DNA; 2763 BP.
AC Q92813;
DT 05-DEC-1995 (first entry)
DE Human MTase.
KW MTase; methyladenosine-phosphatase; malignancy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT exon 254..421
FT /*tag- a 616..720
FT exon b 964..1203
FT /*tag- c 1640..1762
FT exon d 2272..2310
FT /*tag- e
PN WO9518233-A.
PD 06-JUL-1995.
PF 22-DEC-1994; U14920.
PR 29-DEC-1993; US-176855.
PA (CIBA) CIBA GEIGY CORP.
PA (REGC) UNIV CALIFORNIA.
PI Carson DA, Nobori T, Takabayashi K;
DR WPI; 95-246398/32.
PT Detecting methyladenosine phosphatase in mammalian cells - by
PT hybridisation with specific oligonucleotide for detecting malignancy,
PT also new nucleic acid, expression vectors, derived polypeptide(s) and
PT antibodies
PS Claim 8; Page 34-35; 47pp; English.
CC A cosmid gene library constructed from human placenta DNA was
CC screened using a MTase cDNA probe to isolate a human MTase genomic
CC clone (sequence given in Q92813). Absence of MTase from a cell
CC is indicative of malignancy.
SQ Sequence 2763 BP; 778 A; 501 C; 602 G; 805 T;

Query Match 28.3%; Score 240; DB 15; Length 2763;
Best Local Similarity 99.2%; Pred. No. 9.50e-159;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggttctatagaagactgctaagaagctaggactccggtgccaactcaaaaggagacaatgg 1021

QY 570 AGGTTCTTATAGAGACTGCTAAGAGCTAGGACTCCGGTGCCACTCAAAAGGGACAATGG 629
|||||
Db 1022 tcacaatcgaggagacctgcttagctcccgaggagaaagcttcattccgacactggg 1081
|||||
QY 630 TCACAATCGAGGAGGAGCTGTTTACCTCCGGCAGAAAGCTTCATGTTCCGACCTGGG 689
1082 gggcgagatgtatcaacatgaccacagttccagaggtgtgttctgtctaaggagctggaa 1141
|||||
QY 690 GGGCGGATGTTATCAACATGACCACAGTTCAGAGGTGTTCTTGAAGGAGGCTGGA 749
1142 ttgttaccgaagatcgccatgggacagattatgactgtggaagagcagaggaag 1201
|||||
QY 750 TTTGTTACGCAAGTATCGCCATGCGCAGACAGATTATGACTGCTGGAAGAGCAGAGGAG 809
1202 cagt 1205
|||||
QY 810 CAGT 813

RESULT 3
ID Q99202 standard; DNA; 2784 BP.
AC Q99202;
DT 07-MAR-1996 (first entry)
DE Pseudomonas putida methylthioadenosine-phosphorylase DNA.
KW Enzyme; MTase; L-Met-L-deamino-gamma-mercaptopmethane-lyase;
KW chemotherapy; cancer therapy; methionine starvation; ss.
OS Pseudomonas putida.
FH Key Location/Qualifiers
FT CDS 1..2763
FT /*tag- a
PN WO9517908-A1.
PD 06-JUL-1995.
PF 22-DEC-1994; U14919.
PR 29-DEC-1993; US-176413.
PA (REGC) UNIV CALIFORNIA.
PI Carson DA, Nobori T;
DR WPI; 95-246192/32.
PT Selective methionine starvation of methylthioadenosine
PT phosphorylase negative tumour cells - used in chemotherapy of
PT mammalian malignant cells.
PS Disclosure; Page 27-28; 46pp; English.
CC Antigenic peptides may be generated from this sequence encoding
CC methyladenosine-phosphorylase, which are then used to generate
CC antibodies specific for MTase. The produced antibodies may be
CC used in an immunoassay for the detection of MTase.
SQ Sequence 2784 BP; 781 A; 505 C; 598 G; 804 T;

Query Match 28.3%; Score 240; DB 16; Length 2784;
Best Local Similarity 99.2%; Pred. No. 9.50e-159;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 aggttctatagaagactgctaagaagctaggactccggtgccaactcaaaaggagacaatgg 1021
|||||
QY 570 AGGTTCTTATAGAGACTGCTAAGAGCTAGGACTCCGGTGCCACTCAAAAGGGACAATGG 629
tcacaatcgaggagacctgcttagctcccgaggagaaagcttcattccgacactggg 1081
|||||
QY 1022 tcacaatcgaggagacctgcttagctcccgaggagaaagcttcattccgacactggg 1081
|||||
QY 630 TCACAATCGAGGAGGAGCTGTTTACCTCCGGCAGAAAGCTTCATGTTCCGACCTGGG 689
1082 gggcgagatgtatcaacatgaccacagttccagaggtgtgttctgtctaaggagctggaa 1141
|||||
QY 690 GGGCGGATGTTATCAACATGACCACAGTTCAGAGGTGTTCTTGAAGGAGGCTGGA 749
1142 ttgttaccgaagatcgccatgggacagattatgactgtggaagagcagaggaag 1201
|||||
QY 750 TTTGTTACGCAAGTATCGCCATGCGCAGACAGATTATGACTGCTGGAAGAGCAGAGGAG 809
1202 cagt 1205
|||||
QY 810 CAGT 813

```

RESULT 4
ID Q10572 standard; DNA; 1047 BP.
AC Q10572;
DT 09-APR-1991 (first entry)
DE Human Natriuretic Peptide Receptor B.
KW NPRB; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;
KW hyperaldosteronism; glaucoma; guanyl cyclase.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..22
FT /label= signal sequence
FT Protein 12
FT /label= mature NPRB
FT Domain 23..455
FT /label= extracellular domain
FT /note= "binds natriuretic peptides A, B and C]"
FT Domain 456..456
FT /label= transmembrane domain
FT Domain 479..1047
FT /label= cytoplasmic domain
FT /note= "GC and protien kinase activity"
FT Modified-site 24..26
FT /label= N-glycos_site
FT Modified-site 35..37
FT /label= N-glycos_site
FT Modified-site 161..163
FT /label= N-glycos_site
FT Modified-site 195..197
FT /label= N-glycos_site
FT Modified-site 244..246
FT /label= N-glycos_site
FT Modified-site 277..279
FT /label= N-glycos_site
FT Modified-site 349..351
FT /label= N-glycos_site
FT Modified-site 600..602
FT /label= N-glycos_site
FT WO9100292-A.
PN 10-JAN-1991.
PD 22-JUN-1990; U03586.
PF 23-JUN-1989; US-370673.
PA (GETH ) GENENTECH INC.
PI Chang M, Goeddel D, Lowe D;
DR WPI; 91-036711/05.
DR N-PSDB; Q10324.
PT Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PS Claim 3; Fig 1; 49pp; English.
CC The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NPRB, having guanyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NPRB analogues. The protein has a mol wt. of 115 kD (calculated Mr-
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NPRB can
CC also be prep'd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match 8.6%; Score 73; DB 2; Length 1047;
Best Local Similarity 8.3%; Pred. No. 2.34e-33;
Matches 68; Conservative 233; Mismatches 506; Indels 9; Gaps 9;

Db 221 grnvycgnnmhnnnnnnnnnnnnngdvyvnyndvngsnragnrtatgrnwndrtr 280
| : : : : : | : : : : : | : : : : :
Db 131 GGCACCACCACCACCGCCGCTGAAGATTGGAATAATTGGTGGACAGCGCCCTGATGATCCA 190
| : : : : : | : : : : : | : : : : :
Db 281 nnananrnanntvntlyrnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnnagcnydg 340
| : : : : : | : : : : : | : : : : :
QY 191 GAAATTTTAGAAGAGAAGAACTGAAAAATATGTGTGATACTCCATTGTGGCAAGCCATCTGAT 250
| : : : : : | : : : : : | : : : : :
Db 341 nnnyanvnnntnnngtgrndgrrnnvkmgrryhgvtgnv-vmdkndrntdnvnwamgd 399
| : : : : : | : : : : : | : : : : :

```

QY	251	GCCTTAATTTTGGGGAAGATAAAAAATGTTGATTGCATCCCTCTTGCAAGGCATGGAAGG	310
Db	400	ndsgdnnnaahys-ganknwtgrrnnvwkgansdnnncandnddsckltmstnan	458
QY	311	CAGCACACCATCATGCTCCCTTCAAAAGGTCAACTACAGGGCAACATCTGGGCTTTGAAGGAA	370
Db	459	vangtgnntnmgyvsnnnnrknmmknnaasmwrrnrwnnnngnsnrhkgagsrntns	518
QY	371	GAGGGCTGTACACATGTCATAGTAGCACACAGCTTGTGGCTCTTGAGGAGAG-ATT-C	428
Db	519	nrgssysnmtahgkyynnantgtbhkgvvanhvkknrrntrrvnnnkkhmrdrnnnh	578
QY	429	-AGCCCCG-CGATATGTGTCATTATTGATTCAGTTTCATTGACAGACCACTATGAGACTCA	486
Db	579	ntrngacdnnnnncvltncyrgsndnnndnsnndkmnrlysnndrvkymannhnsn	638
QY	487	GTCCTTCTATGATGGAAGTCAATTCCTGTGCGCACAGAGAGTGTGCCATATTCCAATGGCTGA	546
Db	639	nsshsnksncvvdsrnvkntdygnasnrtstannddnnaayakkntaannnsynnn	698
QY	547	GCCGTTTGCCCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAGCTAGACTCCG	606
Db	699	tctgmaadvysngnnnnnanrsngnyngndnsknvnhkvrngnrnyrnsndrtnn	758
QY	607	GTGCCACTCAAAAGGGACATGTGTCACAATCGAGGACCTGCTTTTACCTCCCGGCGAGA	666
Db	759	nnnvnmm-rxwandaanrndngnkngrnrnkngtstnndnnnmnyannnnknvn	817
QY	667	AAGCTTCATGTGTCCGCACCTGGGGGGGATGTATCAACATGACACAGACTTCCAGAGT	726
Db	818	nrtnaynnkrkanannynnnnhsvannnkrgntvnanandsv-tnysndvngtansan	876
QY	727	GGTCTTGCTAAGAGAGGCTGGAATTTGTACCGCAAGTATCGCATGGCAGACAGATTATGA	786
Db	877	stnmnvtnndytcndandndvyk-vnlngdaymvvsngngrngrhannarman	935
QY	787	CTGCTGGAAGAGACGAGAGAGCAGTTTCGTGACCGGGCTTTAAAGACCTGAAAGA	846
Db	936	anndavssnrnrhndnrnrngvhtgvcagvvgnkmmrycngdvtntasrnsngn	995
QY	847	AAACGCTAATAAAGCCAAAGCTTACTGCTCACTACCATACCTCAGATAGGTCACAGA	906
Db	996	anknhvsttkdandnngcnnnnnrgdvnmkqkym	1031
QY	907	ATGGTCAGAAACCCTCCATAACCTGAAGAATATGGC	942
RESULT	5		
ID	Q10572	standard; DNA; 1047 BP.	
AC	Q10572;		
DT	09-APR-1991	(first entry)	
DE	Human Natriuretic Peptide Receptor B.		
KW	NPBR; ANP; BNP; CNP; kidney failure; heart failure; protein kinase;		
KW	hyperaldosteronism; glaucoma; guanyl cyclase.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	Peptide	1..22	
FT	/label= signal sequence		
FT	Protein	12	
FT	/label= mature NPBR		
FT	Domain	23..455	
FT	/label= extracellular domain		
FT	/note= "binds natriuretic peptides A,B and C]"		
FT	Domain	456..456	
FT	/label= transmembrane domain		
FT	Domain	479..1047	
FT	/label= cytoplasmic domain		
FT	/note= "GC and protien kinase activity"		
FT	Modified -site	24..26	
FT	/label= N-glycos_site		
FT	Modified -site	35..37	
FT	/label= N-glycos_site		
FT	Modified -site	161..163	

FT /label- N-glycos_site
FT Modified -site 195..197
FT /label- N-glycos_site
FT Modified -site 244..246
FT /label- N-glycos_site
FT Modified -site 277..279
FT /label- N-glycos_site
FT Modified -site 349..351
FT /label- N-glycos_site
FT Modified -site 600..602
FT /label- N-glycos_site
PN WO9100292-A.
PD 10-JAN-1991.
PF 22-JUN-1990; U03586.
PR 23-JUN-1989; US-370673.
PA (GETH) GENTECH INC.
PI Chang M, Goeddel D, Lowe D;
DR WPI; 91-036711/05.
DR N-PSDB: Q10324.
PS Natriuretic protein receptor B - for diagnosis and treatment of
PT kidney failure, heart failure, hyperaldosteronism, glaucoma etc.
PT Claim 3; Fig 1; 49pp; English.
PS The sequence was derived from the DNA encoding natriuretic peptide
CC receptor B, NMRB, having guanyl cyclase (GC) activity and protein
CC kinase activity. The DNA can be inserted into expression vectors
CC for the prodn. of the protein, opt. after being mutated to produce
CC NMRB analogues. The protein has a mol wt. of 115 kD (calculated Mr
CC 114,952). The protein (or variants) can be used in treatment of
CC natriuretic peptide disorders, and also to isolate peptides using
CC affinity chromatography. Antibodies with affinity for NMRB can
CC also be prepd.
SQ Sequence 1047 BP; 87 A; 15 C; 83 G; 51 T;

Query Match	7.58;	Score 64;	DB 2;	Length 1047;
Best Local Similarity	9.58;	Pred. No. 4.78e-27;		
Matches	77;	Mismatches 501;	Indels 11;	Gaps 11;
		Conservative		

Dp 114 hvrnnntagavaasgnakndhyrltnvrtgnsankngnvnvtlnghnmwtaraannynnda 173
::: | |: : ::::| :| :| :
Cp 931 CAGGTTATGGAGGGTTCGTGACCATTCTCGTGACCCCTATCTGAG-GTATGGTAGTGAGC 874
Dd 174 rtdrrnhyntlngnvnanngsnsvnhvyarlmggnnatlhnrangrnyvcgnmmn 233
: : : :| :| : : :| :| :| :
Cp 873 AGTAAGCTTTGGCTTATTAGCGTTTTCTTCAGGGTCTTTAAGACCCGGTCAACCGAA 814
Db 234 hnannanrnntngdyvnyndvn-gsnragntlratgrnwdrtrnaanaannt 292
| : : : :| :| :| :| :| :| :
Cp 813 ACTGCTTCTCGTGCTCTCTCCAGCAGCATATACTGTGCCCATGGCGATACTTGCGTAA 754
Db 293 vnrvtyrnnnnnnnnnnnnrrarndngyngsnmnneg-chnydgnnnyavnnnt 351
: :| :| :| :| :| :| :| :| :| :| :
Cp 753 CAATTTCCAGCCTCCTTAGACAAGAACACACTGTGAAGTGTGCATGTGTATACATCC 694
Db 352 nnngrtndgnrnvnkmgrrrygvltgnvvmkdndrtdnvnamgdndsgdnnnaahy 411
: :| :| :| :| :| :| :| :| :| :| :
Cp 693 GCCCCCCAGGTGCGGAACATGAAGCTTTCTGCCCCGGAGCTAAACGAGGTCCCTCGATT 634
Db 412 sganknnwltgrnnnwkgaaansdnncanddnscdktnstraanvangtgnmnng 471
: : : :| :| :| :| :| :| :| :| :| :| :
Cp 633 GTGACCATTTGTCCCCTTTGAGTGGCACCC-GGAGTCCTAGCTTCTTAGCACTCTCATATAG 575
Db 472 vssnnnnrknnnkknasmwrnrwnnnnnngnsaryhkagsrnltnsnrgssysnmtah 531
: :| :| :| :| :| :| :| :| :| :| :
Cp 574 AACCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATTGGAATA-TGGCACACTCCTCTGG 516
Db 532 gkynnantgkhkgnvvankhvkkkrnnltrnvnnnkkmrdvnnhnlnrngacndnn 591
| :| :| :| :| :| :| :| :| :| :| :
Cp 515 CACAAGAATGACTTCCATCATATGAAGGACTGAGGTCTCATAGTGTCTGTCAATGAAGT 456
Db 592 ncuvltnyrsgsndnnndsmnd-wmnrysinndnv-kymannhnsnshgsnkssn 649
: :| :| :| :| :| :| :| :| :| :| :
Cp 455 GATCAATATATGACAATATCGCCGGGCTGAATCTCTCTCCCTCAAGAGCACCAAGCTGTGG 396

[illegible]

RESULT 6
ID Q51746 standard; cDNA; 91 BP.
AC Q51746;
DT 31-MAY-1994 (first entry)
DE Oligonucleotide probe MK14-A
KW Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW ss.
OS Synthetic.
PN EP-571911-A.
PD 01-DEC-1993.
PF 24-MAY-1993; 108325.
PR 26-MAY-1992; US-889651.
PA (BECT) BECTON DICKINSON CO.
PI Shank DD, Spears PA;
DR WPI: 93-37884/48.
PT New oligo:nucleotide probes specific for Mycobacteria - used for
PT detection and amplification of Mycobacteria nucleic acid in
PT samples
PS Claim 3; Page 14; 23pp; English.
CC Oligonucleotide probe MK14-A consists of nucleotides 5-95 of MK14
CC (Q51735). It hybridized to all spp. of mycobacteria tested, but
CC cross reacted to a few non-mycobacterial spp. The probe may
CC be useful as an initial screen for mycobacterial infection.
CC See also Q51735-45 and Q51747-59.
SQ Sequence 91 BP; 5 A; 17 C; 15 G; 4 T;

Query Match	Best Local	Similarity	Score	DB	Length
Matches	1, Conservative	2.0%;	44;	9;	91;
		Pred. 1.45e-13;			
		46; Mismatches	3;	Indels	0;
				Gaps	0;

Db	12 svhsyvvhvshhsbvhhvvhvsvvvvvhhvvhvvhhvhyhvyvsc	61
CP	709 CAGGTGATACATCCGCCCCCAGGTGCGGACATGAGCTTTCGCC	660

RESULT	7
ID	Q51746 standard; cDNA; 91 BP.
AC	Q51746;
DT	31-MAY-1994 (first entry)
DE	Oligonucleotide probe MK14-A
KW	Oligonucleotide; DNA probe; mycobacteria; disease diagnosis;
KW	ss.
OS	Synthetic.
PN	EP-571911-A.
PD	01-DEC-1993.
PF	24-MAY-1993; 108325.
PR	26-MAY-1992; US-889651.
PA	(BECT) BECTON DICKINSON CO.
PI	Shank BD, Spears PA;
DR	WPI; 93-378844/48.
PT	New oligo:nucleotide probes specific for Mycobacteria - used for detection and amplification of Mycobacteria nucleic acid in

[illegible]

```

DE Base substituted E.coli beta-galactosidase alpha-fragment.
KW E.coli beta galactosidase alpha-fragment; base substitutions; ss.
OS Escherichia coli.
FH Key Location/Qualifiers
FT misc_feature 19..69
FT /*tag- a
FT /function=multiple cloning site
FT primer_bind 187..204
FT /*tag- b
PN EP-285123-A.
PD 05-MAY-1988.
PE 30-MAR-1988; 105163.
PR 03-APR-1987; US-034819.
PA (SUSO) SUOMEN SOKERI OY.
PI Lehtovaara P, Knowles J, Kolvula A, Bamford J, Reinikainen T;
DR WPI; 88-279927/40.
PT Introducing random point mutations into nucleic acids -
PT by prepn of single stranded template, annealing a primer, elongation,
PT misincorporation, completion of molecules and screening.
PS Disclosure; P; English.
CC Random point mutations were introduced into the alpha fragment of
CC E.coli beta-galactosidase. The wild type sequence was obtained as a
CC single stranded template and an oligonucleotide was hybridised to
CC it to generate a popn of DNA molecules which terminate at all
CC possible nucleotide positions within a specified region. The
CC variable 3' ends generated in this way are used as primers for
CC reverse transcriptase. Nucleotides are misincorporated by the
CC transcriptase and the molecules are completed to forms that can be
CC amplified and then expressed in a suitable host-vector system.
CC The sequence covers all 176 diff base substitutions, most of which
CC occurred singularly in any given mutant.
CC See also P80575.
SQ Sequence 204 BP; 21 A; 47 C; 17 G; 11 T; 108 Others;
SC
Query Match 4.48; Score 37; DB 1; Length 204;
Best Local Similarity 13.18; Pred. No. 4.02e-09;
Matches 13; Conservative 53; Mismatches 31; Indels 2; Gaps 1;
Db 92 hhyrrmbnvrydrnsdaawcyccyrsvkkydcynachdhhyvbybbyvnhnmc 151
::: :::: :::: ::::: |::: :::: ::::: ::::: ::::: ::::: |
QY 828 TCTTAAGACCCCTGAAGAAGAAACGCTATATAAGCCAAAGCTTACTGCTACCATCAC 887
152 ncccbnhvchvbnhbnhrwayrhdarrddvhcvcvh 190
| : :: :::: ::::: |::: ::::: |::: ::::: |
QY 888 CTCAGATAG--GGTCCACAGATGGTCAGAAACCTCCA 924
RESULT 10
ID Q70469 standard; DNA; 114 BP.
AC Q70469;
DT 07-APR-1995 (first entry)
DE Generic DNA sequence to generate a random TSAR peptide library.
KW TSAR; totally synthetic affinity reagent; synthetic; binding domain;
KW effector domain; concatenated heterofunctional protein; linker;
KW direct; rapid; detection; screening; treatment; generic; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT misc_feature 55..60
FT /*tag- a
FT /note="this sequence represents 'Z'; Z can be a
FT sequence of 6,9 or 12 nucleotides (see
FT comments)"
PN WO9418318-A.
PD 18-AUG-1994.
PE 01-FEB-1994; U00977.
PR 01-FEB-1993; US-013416.
PR 30-DEC-1993; US-176500.
PR 31-JAN-1994; US-189331.
PA (UYNC-) UNIV NORTH CAROLINA.
PI Fowlkes DM, Kay BK;
DR WPI; 94-279739/34.
PT Identifying proteins or peptide(s) which bind a ligand - by
PT screening a recombinant vector library expressing fusion proteins

```

MIRAGE (TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Msearch_n n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Feb 10 23:56:55 1997; MasPar time 440.34 Seconds
850.327 Million cell updates/sec

Tabular output not generated.

Title: >US-08-674-311-1
Description: (122-970) from US08674311.seq (2 of 2)
Perfect Score: 849
N.A. Sequence: 122 ATGGCCTCTGGCACCACCAC.....CTGTTTATTACCAAGACAT 970
Comp: TACCGAGAGACCGTGTGTG.....GACAAATAATGCTCTGTA

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 630489 seqs, 220513910 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

EST-STS
1:EST1 2:EST2 3:EST3 4:EST4 5:EST5 6:EST6 7:EST7 8:EST8
9:EST9 10:EST10 11:EST11 12:EST12 13:EST13 14:EST14
15:EST15 16:EST16 17:EST17 18:EST18 19:EST19 20:EST20
21:EST21 22:EST22 23:EST23 24:EST24 25:EST25 26:EST26
27:EST27 28:EST28 29:EST29 30:EST30 31:EST31 32:EST32
33:EST33 34:EST34 35:EST35 36:EST36 37:EST37 38:EST38
39:EST39 40:EST40 41:EST41 42:EST42 43:EST43 44:EST44
45:EST45 46:EST46 47:EST47 48:EST48 49:EST49 50:EST50
51:EST51 52:EST52 53:EST53 54:EST54 55:EST55 56:EST56
57:EST57 58:EST58 59:EST59 60:EST60 61:EST61 62:EST62
63:EST63 64:EST64 65:EST65 66:EST66 67:EST67 68:EST68
69:EST69 70:EST70 71:EST71 72:EST72 73:EST73 74:EST74
75:EST75 76:EST76 77:EST77 78:EST78 79:EST79 80:EST80
81:EST81 82:EST82 83:EST83 84:EST84 85:EST85 86:EST86
87:EST87 88:EST88 89:EST89 90:EST90 91:EST91 92:EST92
93:EST93 94:EST94 95:EST95 96:EST96 97:EST97 98:EST98
99:EST99

Database:

EST-STS-TWO
100:EST100 101:EST101 102:EST102 103:EST103 104:EST104
105:EST105 106:EST106 107:EST107 108:EST108 109:EST109
110:EST110 111:EST111 112:EST112 113:EST113 114:EST114
115:EST115 116:EST116 117:EST117 118:EST118 119:EST119
120:EST120 121:EST121 122:EST122 123:EST123 124:EST124
125:EST125 126:EST126 127:EST127 128:EST128 129:STS1
130:STS2 131:STS3 132:STS4 133:STS5 134:STS6 135:STS7
136:STS8 137:STS9 138:STS10 139:STS11 140:STS12
141:STS13 142:STS14 143:STS15 144:STS16 145:STS17
146:STS18 147:STS19 148:STS20 149:STS21 150:STS22
151:STS23 152:STS24 153:STS25 154:STS26 155:STS27
156:STS28 157:STS29 158:STS30 159:STS31 160:STS32
161:STS33 162:STS34 163:STS35 164:STS36 165:STS37
166:STS38 167:STS39 168:STS40 169:STS41 170:STS42
171:STS43 172:STS44 173:STS45 174:STS46 175:STS47
176:STS48 177:STS49 178:STS50 179:STS51 180:STS52
181:STS53 182:STS54 183:STS55 184:STS56 185:STS57
186:STS58 187:STS59 188:STS60 189:STS61 190:STS62
191:STS63 192:STS64 193:STS65 194:STS66 195:STS67
196:STS68 197:STS69 198:STS70 199:STS71 200:STS72
201:STS73 202:STS74 203:STS75 204:STS76 205:STS77
206:STS78 207:STS79 208:STS80 209:STS81 210:STS82
211:STS83 212:STS84 213:STS85 214:STS86 215:STS87
216:STS88 217:STS89 218:STS90 219:STS91 220:STS92
221:STS93 222:STS94 223:STS95 224:STS96 225:STS97
226:STS98 227:STS99 228:STS100 229:STS101 230:STS102
231:STS103 232:STS104 233:STS105 234:STS106 235:STS107
236:STS108 237:STS109 238:STS110 239:STS111 240:STS112
241:STS113 242:STS114 243:STS115 244:STS116 245:STS117
246:STS118 247:STS119 248:STS120 249:STS121 250:STS122
251:STS123 252:STS124 253:STS125 254:STS126 255:STS127
256:STS128 257:STS129 258:STS130 259:STS131 260:STS132
261:STS133 262:STS134 263:STS135 264:STS136 265:STS137
266:STS138 267:STS139 268:STS140 269:STS141 270:STS142
271:STS143 272:STS144 273:STS145 274:STS146 275:STS147
276:STS148 277:STS149 278:STS150 279:STS151 280:STS152
281:STS153 282:STS154 283:STS155 284:STS156 285:STS157
286:STS158 287:STS159 288:STS160 289:STS161 290:STS162
291:STS163 292:STS164 293:STS165 294:STS166 295:STS167
296:STS168 297:STS169 298:STS170 299:STS171 300:STS172
301:STS173 302:STS174 303:STS175 304:STS176 305:STS177
306:STS178 307:STS179 308:STS180 309:STS181 310:STS182
311:STS183 312:STS184 313:STS185 314:STS186 315:STS187
316:STS188 317:STS189 318:STS190 319:STS191 320:STS192
321:STS193 322:STS194 323:STS195 324:STS196 325:STS197
326:STS198 327:STS199 328:STS200 329:STS201 330:STS202
331:STS203 332:STS204 333:STS205 334:STS206 335:STS207
336:STS208 337:STS209 338:STS210 339:STS211 340:STS212
341:STS213 342:STS214 343:STS215 344:STS216 345:STS217
346:STS218 347:STS219 348:STS220 349:STS221 350:STS222
351:STS223 352:STS224 353:STS225 354:STS226 355:STS227
356:STS228 357:STS229 358:STS230 359:STS231 360:STS232
361:STS233 362:STS234 363:STS235 364:STS236 365:STS237
366:STS238 367:STS239 368:STS240 369:STS241 370:STS242
371:STS243 372:STS244 373:STS245 374:STS246 375:STS247
376:STS248 377:STS249 378:STS250 379:STS251 380:STS252
381:STS253 382:STS254 383:STS255 384:STS256 385:STS257
386:STS258 387:STS259 388:STS260 389:STS261 390:STS262
391:STS263 392:STS264 393:STS265 394:STS266 395:STS267
396:STS268 397:STS269 398:STS270 399:STS271 400:STS272
401:STS273 402:STS274 403:STS275 404:STS276 405:STS277
406:STS278 407:STS279 408:STS280 409:STS281 410:STS282
411:STS283 412:STS284 413:STS285 414:STS286 415:STS287
416:STS288 417:STS289 418:STS290 419:STS291 420:STS292
421:STS293 422:STS294 423:STS295 424:STS296 425:STS297
426:STS298 427:STS299 428:STS300 429:STS301 430:STS302
431:STS303 432:STS304 433:STS305 434:STS306 435:STS307
436:STS308 437:STS309 438:STS310 439:STS311 440:STS312
441:STS313 442:STS314 443:STS315 444:STS316 445:STS317
446:STS318 447:STS319 448:STS320 449:STS321 450:STS322
451:STS323 452:STS324 453:STS325 454:STS326 455:STS327
456:STS328 457:STS329 458:STS330 459:STS331 460:STS332
461:STS333 462:STS334 463:STS335 464:STS336 465:STS337
466:STS338 467:STS339 468:STS340 469:STS341 470:STS342
471:STS343 472:STS344 473:STS345 474:STS346 475:STS347
476:STS348 477:STS349 478:STS350 479:STS351 480:STS352
481:STS353 482:STS354 483:STS355 484:STS356 485:STS357
486:STS358 487:STS359 488:STS360 489:STS361 490:STS362
491:STS363 492:STS364 493:STS365 494:STS366 495:STS367
496:STS368 497:STS369 498:STS370 499:STS371 500:STS372
501:STS373 502:STS374 503:STS375 504:STS376 505:STS377
506:STS378 507:STS379 508:STS380 509:STS381 510:STS382
511:STS383 512:STS384 513:STS385 514:STS386 515:STS387
516:STS388 517:STS389 518:STS390 519:STS391 520:STS392
521:STS393 522:STS394 523:STS395 524:STS396 525:STS397
526:STS398 527:STS399 528:STS400 529:STS401 530:STS402
531:STS403 532:STS404 533:STS405 534:STS406 535:STS407
536:STS408 537:STS409 538:STS410 539:STS411 540:STS412
541:STS413 542:STS414 543:STS415 544:STS416 545:STS417
546:STS418 547:STS419 548:STS420 549:STS421 550:STS422
551:STS423 552:STS424 553:STS425 554:STS426 555:STS427
556:STS428 557:STS429 558:STS430 559:STS431 560:STS432
561:STS433 562:STS434 563:STS435 564:STS436 565:STS437
566:STS438 567:STS439 568:STS440 569:STS441 570:STS442
571:STS443 572:STS444 573:STS445 574:STS446 575:STS447
576:STS448 577:STS449 578:STS450 579:STS451 580:STS452
581:STS453 582:STS454 583:STS455 584:STS456 585:STS457
586:STS458 587:STS459 588:STS460 589:STS461 590:STS462
591:STS463 592:STS464 593:STS465 594:STS466 595:STS467
596:STS468 597:STS469 598:STS470 599:STS471 600:STS472
601:STS473 602:STS474 603:STS475 604:STS476 605:STS477
606:STS478 607:STS479 608:STS480 609:STS481 610:STS482
611:STS483 612:STS484 613:STS485 614:STS486 615:STS487
616:STS488 617:STS489 618:STS490 619:STS491 620:STS492
621:STS493 622:STS494 623:STS495 624:STS496 625:STS497
626:STS498 627:STS499 628:STS500 629:STS501 630:STS502
631:STS503 632:STS504 633:STS505 634:STS506 635:STS507
636:STS508 637:STS509 638:STS510 639:STS511 640:STS512
641:STS513 642:STS514 643:STS515 644:STS516 645:STS517
646:STS518 647:STS519 648:STS520 649:STS521 650:STS522
651:STS523 652:STS524 653:STS525 654:STS526 655:STS527
656:STS528 657:STS529 658:STS530 659:STS531 660:STS532
661:STS533 662:STS534 663:STS535 664:STS536 665:STS537
666:STS538 667:STS539 668:STS540 669:STS541 670:STS542
671:STS543 672:STS544 673:STS545 674:STS546 675:STS547
676:STS548 677:STS549 678:STS550 679:STS551 680:STS552
681:STS553 682:STS554 683:STS555 684:STS556 685:STS557
686:STS558 687:STS559 688:STS560 689:STS561 690:STS562
691:STS563 692:STS564 693:STS565 694:STS566 695:STS567
696:STS568 697:STS569 698:STS570 699:STS571 700:STS572
701:STS573 702:STS574 703:STS575 704:STS576 705:STS577
706:STS578 707:STS579 708:STS580 709:STS581 710:STS582
711:STS583 712:STS584 713:STS585 714:STS586 715:STS587
716:STS588 717:STS589 718:STS590 719:STS591 720:STS592
721:STS593 722:STS594 723:STS595 724:STS596 725:STS597
726:STS598 727:STS599 728:STS600 729:STS601 730:STS602
731:STS603 732:STS604 733:STS605 734:STS606 735:STS607
736:STS608 737:STS609 738:STS610 739:STS611 740:STS612
741:STS613 742:STS614 743:STS615 744:STS616 745:STS617
746:STS618 747:STS619 748:STS620 749:STS621 750:STS622
751:STS623 752:STS624 753:STS625 754:STS626 755:STS627
756:STS628 757:STS629 758:STS630 759:STS631 760:STS632
761:STS633 762:STS634 763:STS635 764:STS636 765:STS637
766:STS638 767:STS639 768:STS640 769:STS641 770:STS642
771:STS643 772:STS644 773:STS645 774:STS646 775:STS647
776:STS648 777:STS649 778:STS650 779:STS651 780:STS652
781:STS653 782:STS654 783:STS655 784:STS656 785:STS657
786:STS658 787:STS659 788:STS660 789:STS661 790:STS662
791:STS663 792:STS664 793:STS665 794:STS666 795:STS667
796:STS668 797:STS669 798:STS670 799:STS671 800:STS672
801:STS673 802:STS674 803:STS675 804:STS676 805:STS677
806:STS678 807:STS679 808:STS680 809:STS681 810:STS682
811:STS683 812:STS684 813:STS685 814:STS686 815:STS687
816:STS688 817:STS689 818:STS690 819:STS691 820:STS692
821:STS693 822:STS694 823:STS695 824:STS696 825:STS697
826:STS698 827:STS699 828:STS700 829:STS701 830:STS702
831:STS703 832:STS704 833:STS705 834:STS706 835:STS707
836:STS708 837:STS709 838:STS710 839:STS711 840:STS712
841:STS713 842:STS714 843:STS715 844:STS716 845:STS717
846:STS718 847:STS719 848:STS720 849:STS721 850:STS722
851:STS723 852:STS724 853:STS725 854:STS726 855:STS727
856:STS728 857:STS729 858:STS730 859:STS731 860:STS732
861:STS733 862:STS734 863:STS735 864:STS736 865:STS737
866:STS738 867:STS739 868:STS740 869:STS741 870:STS742
871:STS743 872:STS744 873:STS745 874:STS746 875:STS747
876:STS748 877:STS749 878:STS750 879:STS751 880:STS752
881:STS753 882:STS754 883:STS755 884:STS756 885:STS757
886:STS758 887:STS759 888:STS760 889:STS761 890:STS762
891:STS763 892:STS764 893:STS765 894:STS766 895:STS767
896:STS768 897:STS769 898:STS770 899:STS771 900:STS772
901:STS773 902:STS774 903:STS775 904:STS776 905:STS777
906:STS778 907:STS779 908:STS780 909:STS781 910:STS782
911:STS783 912:STS784 913:STS785 914:STS786 915:STS787
916:STS788 917:STS789 918:STS790 919:STS791 920:STS792
921:STS793 922:STS794 923:STS795 924:STS796 925:STS797
926:STS798 927:STS799 928:STS800 929:STS801 930:STS802
931:STS803 932:STS804 933:STS805 934:STS806 935:STS807
936:STS808 937:STS809 938:STS810 939:STS811 940:STS812
941:STS813 942:STS814 943:STS815 944:STS816 945:STS817
946:STS818 947:STS819 948:STS820 949:STS821 950:STS822
951:STS823 952:STS824 953:STS825 954:STS826 955:STS827
956:STS828 957:STS829 958:STS830 959:STS831 960:STS832
961:STS833 962:STS834 963:STS835 964:STS836 965:STS837
966:STS838 967:STS839 968:STS840 969:STS841 970:STS842
971:STS843 972:STS844 973:STS845 974:STS846 975:STS847
976:STS848 977:STS849 978:STS850 979:STS851 980:STS852
981:STS853 982:STS854 983:STS855 984:STS856 985:STS857
986:STS858 987:STS859 988:STS860 989:STS861 990:STS862
991:STS863 992:STS864 993:STS865 994:STS866 995:STS867
996:STS868 997:STS869 998:STS870 999:STS871 1000:STS872
1001:STS873 1002:STS874 1003:STS875 1004:STS876 1005:STS877
1006:STS878 1007:STS879 1008:STS880 1009:STS881 1010:STS882
1011:STS883 1012:STS884 1013:STS885 1014:STS886 1015:STS887
1016:STS888 1017:STS889 1018:STS890 1019:STS891 1020:STS892
1021:STS893 1022:STS894 1023:STS895 1024:STS896 1025:STS897
1026:STS898 1027:STS899 1028:STS900 1029:STS901 1030:STS902
1031:STS903 1032:STS904 1033:STS905 1034:STS906 1035:STS907
1036:STS908 1037:STS909 1038:STS910 1039:STS911 1040:STS912
1041:STS913 1042:STS914 1043:STS915 1044:STS916 1045:STS917
1046:STS918 1047:STS919 1048:STS920 1049:STS921 1050:STS922
1051:STS923 1052:STS924 1053:STS925 1054:STS926 1055:STS927
1056:STS928 1057:STS929 1058:STS930 1059:STS931 1060:STS932
1061:STS933 1062:STS934 1063:STS935 1064:STS936 1065:STS937
1066:STS938 1067:STS939 1068:STS940 1069:STS941 1070:STS942
1071:STS943 1072:STS944 1073:STS945 1074:STS946 1075:STS947
1076:STS948 1077:STS949 1078:STS950 1079:STS951 1080:STS952
1081:STS953 1082:STS954 1083:STS955 1084:STS956 1085:STS957
1086:STS958 1087:STS959 1088:STS960 1089:STS961 1090:STS962
1091:STS963 1092:STS964 1093:STS965 1094:STS966 1095:STS967
1096:STS968 1097:STS969 1098:STS970 1099:STS971 1100:STS972
1101:STS973 1102:STS974 1103:STS975 1104:STS976 1105:STS977
1106:STS978 1107:STS979 1108:STS980 1109:STS981 1110:STS982
1111:STS983 1112:STS984 1113:STS985 1114:STS986 1115:STS987
1116:STS988 1117:STS989 1118:STS990 1119:STS991 1120:STS992
1121:STS993 1122:STS994 1123:STS995 1124:STS996 1125:STS997
1126:STS998 1127:STS999 1128:STS1000 1129:STS1001 1130:STS1002
1131:STS1003 1132:STS1004 1133:STS1005 1134:STS1006 1135:STS1007
1136:STS1008 1137:STS1009 1138:STS1010 1139:STS1011 1140:STS1012
1141:STS1013 1142:STS1014 1143:STS1015 1144:STS1016 1145:STS1017
1146:STS1018 1147:STS1019 1148:STS1020 1149:STS1021 1150:STS1022
1151:STS1023 1152:STS1024 1153:STS1025 1154:STS1026 1155:STS1027
1156:STS1028 1157:STS1029 1158:STS1030 1159:STS1031 1160:STS1032
1161:STS1033 1162:STS1034 1163:STS1035 1164:STS1036 1165:STS1037
1166:STS1038 1167:STS1039 1168:STS1040 1169:STS1041 1170:STS1042
1171:STS1043 1172:STS1044 1173:STS1045 1174:STS1046 1175:STS1047
1176:STS1048 1177:STS1049 1178:STS1050 1179:STS1051 1180:STS1052
1181:STS1053 1182:STS1054 1183:STS1055 1184:STS1056 1185:STS1057
1186:STS1058 1187:STS1059 1188:STS1060 1189:STS1061 1190:STS1062
1191:STS1063 1192:STS1064 1193:STS1065 1194:STS1066 1195:STS1067
1196:STS1068 1197:STS1069 1198:STS1070 1199:STS1071 1200:STS1072
1201:STS1073 1202:STS1074 1203:STS1075 1204:STS1076 1205:STS1077
1206:STS1078 1207:STS1079 1208:STS1080 1209:STS1081 1210:STS1082
1211:STS1083 1212:STS1084 1213:STS1085 1214:STS1086 1215:STS1087
1216:STS1088 1217:STS1089 1218:STS1090 1219:STS1091 1220:STS1092
1221:STS1093 1222:STS1094 1223:STS1095 1224:STS1096 1225:STS1097
1226:STS1098 1227:STS1099 1228:STS1100 1229:STS1101 1230:STS1102
1231:STS1103 1232:STS1104 1233:STS1105 1234:STS1106 1235:STS1107
1236:STS1108 1237:STS1109 1238:STS1110 1239:STS1111 1240:STS1112
1241:STS1113 1242:STS1114 1243:STS1115 124


```

AUTHORS      Konishi,K., Morishima,Y.-I., Ueda,E., Nonomura,K., Kibe,S.,
JOURNAL      Yamanishi,K. and Yasuno,H.
REFERENCE    Unpublished (1994)
AUTHORS      2 (sites)
TITLE        Konishi,K., Morishima,Y., Ueda,E., Kibe,Y., Nonomura,K.,
              Yamanishi,K. and Yasuno,H.
JOURNAL      Cataloging of the genes expressed in human keratinocytes: analysis
MEDLINE      of 607 randomly isolated cDNA sequences
COMMENT      Blochem. Biophys. Res. Commun. 202 (2), 976-983 (1994)
              94324994
              Submitted (11-Mar-1994) to DBJ by:
              Kiyofumi Yamanishi
              Department of Dermatology
              Kyoto Prefectural University of Medicine
              Kamigyo-ku
              Kyoto, Kyoto, 602
              Japan
              Phone: 75-251-5587
              Fax : 75-251-5587
              PROJECT='human_epidermal_keratinocyte_ESTs'.
              location/Qualifiers
              1..168
              /organism="Homo sapiens"
              /cell_line="Primary culture"
              /cell_type="Keratinocyte"
              /sequenced_mol="cDNA to mRNA"
              /tissue_type="Epidermis"

FEATURES
source
BASE COUNT   43 a      26 c      37 g      62 t
ORIGIN
Query Match          7.9%; Score 67; DB 50; Length 168;
Best Local Similarity 93.3%; Pred. No. 1.79e-90;
Matches 83; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Db 19 agatttgaataat-ggttgaacagcgccttgatgatccagaataattagaaggaagaact 77
    |||
Db 153 AGATTGGAATATATGTTGGTGAACAGAGCCCTGGATGATCCAGAAATTTTGAAGGAAGAACTG 212
    |||
Db 78 ttaaatatgtgatcac-ccatttggcaag 105
    |||
QY 213 AAAAATATGTGATACTCCATTGGCAAG 241

RESULT 2 G17388 400 bp DNA STS 05-MAR-1996
LOCUS      human STS SHGC-14428 clone pg-7320.
DEFINITION G17388
ACCESSION  G1214814
NID
KEYWORDS   STS sequence; primer; sequence tagged site.
SOURCE     human Plasmid clones, generated from a lymphoblastoid cell line
           from a human male. Localized to human chromosome 9 by analysis on
           the NIGMS Human/Rodent Somatic Cell Hybrid Panel #1, Coriell
           Institute for Medical Research, Camden, NJ 08103.
ORGANISM   Homo sapiens
            Eukaryota; Eukaryota; Eukaryota; Metazoa; Chordata;
            Vertebrata; Gnathostomata; Osteichthyes; Sarcopterygii; Choanata;
            Tetrapoda; Amniota; Mammalia; Theria; Eutheria; Archonta; Primates;
            Catarrhini; Homnidae; Homo.
REFERENCE  1 (bases 1 to 400)
AUTHORS    Myers,R.M.
JOURNAL    Unpublished (1996)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myerseshgc.stanford.edu

Primer A: TGTGATATCCATTGGCA
Primer B: TCACAAATCAAGCCAGCAAG
STS size: 90

```

```

PCR Profile:
Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30
Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
Primer: each 1 uM
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer:
MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

FEATURES
source
Chromosome 9.
Location/Qualifiers
1..400
/organism="Homo sapiens"
/note="human"
61..150
/map="9"
61..80
/map="9"
complement(131..150)
primer_bind
primer_bind
/map="9"
BASE COUNT 95 a 74 c 83 g 101 t 47 others
ORIGIN

Query Match 5.8%; Score 49; DB 134; Length 400;
Best Local Similarity 98.2%; Pred. No. 9.95e-53;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 28 atccgagaattttagaaggaaga-ctgaaatatgtgatccattgccaag 82
|||||
186 ATCCAGAAATTTAGAGGAGAGAGAAATATGTGATCTCATTTGGCAAG 241

RESULT 3
ID HS388293 standard; DNA; STS; 400 BP.
AC G17388;
DT 07-MAR-1996 (Rel. 47, Created)
DE 07-MAR-1996 (Rel. 47, Last updated, Version 1)
DE human STS SHGC-14428 clone pg-7320.
OS primer; sequence tagged site; STS sequence.
OS Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-400
RA Myers R.M.;
RT ;
RL Unpublished.
CC Contact: Richard M. Myers Stanford Human Genome Center (SHGC)
CC Stanford University School of Medicine Department of Genetics,
CC M-344, Stanford, CA 94305, USA Tel: 4157259687 Fax: 4157259689
CC Email: myers@shgc.stanford.edu Primer A: TGTGATACTCTCATTTGGCA
CC Primer B: TCACAAATCAGCCAGCAAG STS size: 90 PCR Profile: Initial
CC incubation: 94 degrees C for 90 seconds Denaturation: 94 degrees C
CC for 15 seconds Annealing: 62 degrees C for 23 seconds
CC Polymerization: 72 degrees C for 30 seconds PCR Cycles: 30 Thermal
CC Cycler: Perkin Elmer 9600 Protocol: Template: 25 ng Primer: each 1
CC uM dNTPs: each 200 uM Taq Polymerase: 0.05 units/ul Total Vol: 10
CC ul Buffer: MgCl2: 2.5 mM KCl: 50 mM Tris-HCl: 20 mM pH: 8.3
CC Chromosome 9. NCBI g1: 1214814
FH location/Qualifiers
FH

```

FT source 1..400
/organism="Homo sapiens"
FT STS /note="human"
FT primer_bind 61..150
FT primer_bind 61..80
FT primer_bind 61..80
FT primer_bind /map="9"
FT primer_bind complement(131..150)
FT primer_bind /map="9"
SQ Sequence 400 BP; 95 A; 74 C; 83 G; 101 T; 47 other;

Query Match 5 8%; Score 49; DB 173; Length 400;
Best Local Similarity 98.2%; Pred. No. 9.95e-53;
Matches 55; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Db 28 atccagaattttagaaggaga-ctgaaaaatatgtgatactccattggcaag 82
186 ATCCAGAAATTTAGAGAGAGAACTGAAAAATATGTGATACCTCCATTGGCAAG 241

RESULT 4
LOCUS CELK008E4F 447 bp mRNA EST 07-AUG-1994
DEFINITION C.elegans CDNA clone yk8e4 : 5' end, single read.
ACCESSION D37705
NID 9526087
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain CB1489 him-8(e1489)) (library: Yuji Kohara unpublished CDNA) Hermaphrodite, male varied whole animal CDNA to mRNA.

ORGANISM Caenorhabditis elegans
Eukaryota; Animalia; Eumetazoa; Nematoda; Secernentea; Rhabditiida;
Rhabditina; Rhabditoidea; Rhabditidae.
REFERENCE 1 (bases 1 to 447)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and Tabara,H.
TITLE Toward an expression map of the C.elegans genome(5'end)
JOURNAL Unpublished (1994)
COMMENT Submitted (1-Jul-1994) to DDBJ by: Yuji Kohara
National Institute of Genetics
Gene Library Lab
Yata 1111, Mishima
Shizuoka 411
Phone: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.

FEATURES
source location/Qualifiers
1..447
/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/dev_stage="varied"
/sequenced_mol="CDNA to mRNA"
/sex="Hermaphrodite, male"
/tissue_type="whole animal"
/clone_lib="Yuji Kohara unpublished CDNA"
BASE COUNT 153 a 114 c 77 g 103 t
ORIGIN

Query Match 2 8%; Score 24; DB 2; Length 447;
Best Local Similarity 80.0%; Pred. No. 8.11e-08;
Matches 32; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Db 113 atgttgataatccattggcaatccatctcacgacatgat 152
219 ATGTGATACTCCATTTGGCAAGCCATCTGATGCCCTTAAT 258

RESULT 5
LOCUS ATTS2467 321 bp RNA EST 14-MAR-1994
DEFINITION A. thaliana transcribed sequence; clone TAI332; 3' end; Similar to ribosomal protein L7a ; Oryza sativa.
ACCESSION Z30721
NID 9460354
KEYWORDS expressed sequence tag; partial CDNA sequence.

SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; mitochondrial eukaryotes; Chlorophyta/Embryophyta group; Charophyta/Embryophyta group; Embryophyta; Magnoliophyta; Magnoliopsida; Dillenidae; Caprales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 321)
AUTHORS Phillips,G. and Gigot,C.
TITLE Direct Submission
JOURNAL Submitted (13-MAR-1994) to the EMBL/GenBank/DBJ databases. CNRS, GDR-1003 ACS, INRA, laboratoire de Biologie Molculaire, BP 27, 31326 Castanet-Tolosan cedex, France.
E-mail:gdr-svpetoulouse.inra.fr. On behalf of: Laboratoire de Biologie Molculaire des Plantes - CNRS, Gigot Claude / L512, 12 Rue du General Zimmer, 67084 Strasbourg Cedex, France.
E-mail:philippe@scilla.u-strasbg.fr
2 (bases 1 to 321)
CNRS.

REFERENCE The Arabidopsis thaliana transcribed genome: the GDR CDNA program
AUTHORS Unpublished
TITLE Cloning vector: lambda ZAPII;
JOURNAL Physiological condition: cycling cells.
COMMENT full automatic.
similarity detected by blastx against GB entry D12631.

FEATURES
source location/Qualifiers
1..321
/organism="Arabidopsis thaliana"
/clone="TAI332"
/tissue_type="cell suspension culture of ecotype columbia"
/clone_lib="AC16H"
complement(188..>321)
/codon_start=3
/product="ribosomal protein"
/db_xref="PID:9460355"
/translation="LTSPTRYEDDRKKWGGMGMSQAKTKAKERYIAKEAAQRMN"

CDS

BASE COUNT 99 a 68 c 51 g 103 t
ORIGIN

Query Match 2 7%; Score 23; DB 1; Length 321;
Best Local Similarity 78.0%; Pred. No. 2.15e-06;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 205 cagcctccttagcaataaccctctccttagcttgcttg 245
746 CAGCCTCCTTAGCAAGAACCACTCTGGAAGTGTGTCATG 706

RESULT 6
LOCUS CELK002F5F 341 bp mRNA EST 06-AUG-1994
DEFINITION C.elegans CDNA clone yk2f5 : 5' end, single read.
ACCESSION D36187
NID 9525402
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain CB1489 him-8(e1489)) (library: Yuji Kohara unpublished CDNA) Hermaphrodite, male varied whole animal CDNA to mRNA.

ORGANISM Caenorhabditis elegans
Eukaryota; Animalia; Eumetazoa; Nematoda; Secernentea; Rhabditiida;
Rhabditina; Rhabditoidea; Rhabditidae.
REFERENCE 1 (bases 1 to 341)
AUTHORS Kohara,Y., Mitsuki,H., Nishigaki,A., Motohashi,T., Sugimoto,A. and Tabara,H.
TITLE Toward an expression map of the C.elegans genome(5'end)
JOURNAL Unpublished (1994)
COMMENT Submitted (1-Jul-1994) to DDBJ by: Yuji Kohara
National Institute of Genetics
Gene Library Lab
Yata 1111, Mishima
Shizuoka 411
Phone: 0559-75-0771
Fax: 0559-75-6240
Email: ykohara@dbj.nig.ac.jp.

FEATURES
source location/Qualifiers
1..341

/organism="Caenorhabditis elegans"
/strain="CB1489 him-8(e1489)"
/dev_stage="varied"
/sequenced_mol="cDNA to mRNA"
/sex="Hermaphrodite, male"
/tissue_type="whole animal"
/clone_lib="Yuji Kohara unpublished cDNA"
111 a 85 c 52 g 89 t 4 others

BASE COUNT
ORIGIN

Query Match 2.7%; Score 23; DB 2; Length 341;
Best Local Similarity 89.7%; Pred. No. 2,15e-06;
Matches 26; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 306 atgttataatccatttggcaatccatct 334
|||||
Qy 219 ATGTGGATACTCCATTGGCAAGCCATCT 247

RESULT 7
LOCUS T22905 435 bp mRNA EST 28-MAR-1995
DEFINITION 4913 Arabidopsis thaliana cDNA clone 108B10T7.
ACCESSION T22905
NID 9504346
KEYWORDS EST.
SOURCE thale cress clone-108B10T7 library-lambda-PRL2 strain-var columbia
vector-lambda Zip-Lox Rsite1-Sal Rsite2-Not Lambda PRL2 is a cDNA
library derived from equal quantities of 4 pools of mRNA. The mRNA
sources were 1) 7 day germinated etiolated seedlings; 2) tissue
culture grown roots; 3) staged plants half with 24 hour light
cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants
as 3 but aerial tissue (stems, flowers and siliques). The vector is
BRL's lambda Zip-Lox. The cDNA inserts were directionally cloned
with Sal-Not arms using oligo dt primed cDNA.

ORGANISM Arabidopsis thaliana
Eucaryotae; Embryophyta; Magnoliophyta; Magnoliopsida; Capparales;
Brassicaceae; Arabidopsids.
1 (bases 1 to 435)
Newman, T., de Bruijn, F.J., Green, P., Keegstra, K., Kende, H.,
McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomashow, M.,
Retzel, E. and Somerville, C.
Genes galore: a summary of methods for accessing results from
large-scale partial sequencing of anonymous Arabidopsis cDNA clones
Plant Physiol. 106, 1241-1255 (1994)

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313tcn@lhm.cl.msu.edu.

FEATURES
source
1..435
/organism="Arabidopsis thaliana"
/clone="108B10T7"
/strain="var columbia"
/note="thale cress"

BASE COUNT 114 a 86 c 112 g 115 t 8 others

ORIGIN

Query Match 2.7%; Score 23; DB 107; Length 435;
Best Local Similarity 78.0%; Pred. No. 2,15e-06;
Matches 32; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Db 294 caagaccaaagctaaggagaggtatttctaaggagctg 334
|||||
Qy 706 CATGACCACAGTTCACAGAGGTGTTCTTGTAAAGGAGCTG 746

RESULT 8
LOCUS CELK062D9R 332 bp mRNA EST 09-DEC-1995

DEFINITION C.elegans cDNA clone yk62d9 : 3' end, single read.
ACCESSION D71125
NID 91110834
KEYWORDS EST(expressed sequence tag).
SOURCE Caenorhabditis elegans (strain N2,) (library: Yuji Kohara
unpublished cDNA) Hermaphrodite embryo embryo cDNA to mRNA.
ORGANISM Caenorhabditis elegans
Eucaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Secernentea; Rhabditia; Rhabditida; Rhabditina; Rhabditidae;
Rhabditidae; Caenorhabditis.
1 (bases 1 to 332)
Kohara, Y., Motohashi, T., Tabara, H., Sugimoto, A., Watanabe, H. and
Nishigaki, A.
Toward an expression map of the C.elegans genome
Unpublished (1995)
COMMENT Submitted (23-Aug-1995) to DDBJ by:
Yuji Kohara
Gene Library Lab.
National Institute of Genetics
Yata 1111, Mishima Shizuoka
411 Japan
Phone: 0559-81-6854
Fax : 0559-81-6855
Email: ykohara@dbj.nig.ac.jp.

FEATURES
source
1..332
/organism="Caenorhabditis elegans"
/strain="N2"
/dev_stage="embryo"
/sequenced_mol="cDNA to mRNA"
/sex="Hermaphrodite"
/tissue_type="embryo"
/clone_lib="Yuji Kohara unpublished cDNA"
130 a 46 c 75 g 81 t

BASE COUNT
ORIGIN

Query Match 2.6%; Score 22; DB 4; Length 332;
Best Local Similarity 92.3%; Pred. No. 5,08e-05;
Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 193 aataaacagaaaatggcgcatatt 218
|||||
Cp 960 AATAAACAGAAACTGGGCCATATT 935

RESULT 9
LOCUS W13162 981 bp mRNA EST 26-APR-1996
DEFINITION ma74d12.r1 Soares mouse p33NF19.5 Mus musculus cDNA 5' similar to
SW:CDK4_MOUSE P30285 CELL DIVISION PROTEIN KINASE 4 ;
W13162
NID 91287199
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eucaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
1 (bases 1 to 981)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.,
Gelsel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M.,
Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B.,
Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and
Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished (1996)

TITLE
JOURNAL
COMMENT
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information.
Seq primer: ETPRimer
High quality sequence stop: 258.

NCBI gi: 1287199

FEATURES
Location/Qualifiers
1..981

Source
/organism="Mus musculus"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTTACCAATCTGAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo."
/clone_lib="Soares mouse p3NMF19.5"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
<1..>981
MRNA
BASE COUNT 203 a 227 c 304 g 236 t 11 others
ORIGIN

Query Match 2.6%; Score 22; DB 150; Length 981;
Best Local Similarity 86.5%; Pred. No. 5.08e-05;
Matches 32; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 328 cacagttcgtgagtgac-ctgttaagagagctgga 363
||||| ||||| ||| ||||| |||||
OY 712 CACAGTTCAGAGGTGTTCTTGTCTAAGAGAGCTGGA 748

RESULT 10
ID NM1627 standard; RNA; EST; 981 BP.
AC W13162;
DT 30-APR-1996 (Rel. 47, Created)
DT 30-APR-1996 (Rel. 47, Last updated, Version 1)
DE ma74d12.r1 Soares mouse p3NMF19.5 Mus musculus cDNA 5' similar to
DE SW:CDK4_MOUSE P30285 CELL DIVISION PROTEIN KINASE 4 ;.
KW EST.
OS Mus musculus (mouse)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Rodentia; Myomorpha; Muridae; Murinae.
RN [1]
RP 1-981
RA Maria M., Hillier L., Allen M., Bowles M., Dietrich N.,
RA Dubuque T., Geisel S., Kucaba T., Lacy M., Le M., Martin J.,
RA Morris M., Schellenberg K., Steptoe M., Tan F., Underwood K.,
RA Moore B., Theising B., Wylie T., Lennon G., Soares B., Wilson R.,
RA Waterston R.;
RT "The WashU-HMI Mouse EST Project";
RL Unpublished.
CC Contact: Maira M/Mouse EST Project WashU-HMI Mouse EST Project
CC Washington University School of Medicine 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: mouseest@wustl.edu This clone is available
CC royalty-free through LNL ; contact the IMAGE Consortium
CC (info@image.lnl.gov) for further information. Seq primer: ETPRimer
CC High quality sequence stop: 258. NCBI gi: 1287199
FH Key
FH Location/Qualifiers
FT 1..981
FT source

/organism="Mus musculus"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5', TGTTACCAATCTGAGTGGAGCGCCGCAATTTTCTTTTCTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector

FT (Pharmacia). Library went through one round of
FT normalization to a Cot = 5. Library constructed by Bento
FT Soares and M.Fatima Bonaldo."
FT /clone_lib="Soares mouse p3NMF19.5"
FT /dev_stage="19 weeks"
FT /lab_host="DH10B (ampicillin resistant)"
FT <1..>981
FT MRNA
FT Sequence 981 BP; 203 A; 227 C; 304 G; 236 T; 11 other;
SQ

Query Match 2.6%; Score 22; DB 169; Length 981;
Best Local Similarity 86.5%; Pred. No. 5.08e-05;
Matches 32; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

Db 328 cacagttcgtgagtgac-ctgttaagagagctgga 363
||||| ||||| ||| ||||| |||||
OY 712 CACAGTTCAGAGGTGTTCTTGTCTAAGAGAGCTGGA 748

RESULT 11
ID HSW04230 standard; RNA; EST; 202 BP.
AC W04230;
DT 30-APR-1996 (Rel. 47, Created)
DT 30-APR-1996 (Rel. 47, Last updated, Version 1)
DE za43e02.r1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
DE 295322 5'.
KW EST.
KW Homo sapiens (human)
OC Eukaryota; Animalia; Metazoa; Chordata; Vertebrata; Mammalia;
OC Theria; Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
RN [1]
RP 1-202
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Maira M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevasakis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
CC Contact: Wilson RK WashU-Merck EST Project Washington University
CC School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis,
CC MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email:
CC est@wustl.edu This clone is available royalty-free through
CC LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for
CC further information. Seq primer: mob.REGA+ET High quality sequence
CC stop: 176. NCBI gi: 1276338
FH Key
FH Location/Qualifiers
FH source

1..202
/organism="Homo sapiens"
/note="Organ: Liver and spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dt) primer
[5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTT 3'] ,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library wen
t
through one round of normalization. Library constructed by
Bento Soares and M.Fatima Bonaldo."
/clone="295322"
/clone_lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
<1..>202
MRNA
SQ Sequence 202 BP; 80 A; 33 C; 43 G; 44 T; 2 other;

Query Match 2.5%; Score 21; DB 168; Length 202;
Best Local Similarity 88.9%; Pred. No. 1.06e-03;
Matches 24; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Db 19 gaccataactgagtgatagtgagc 45
||||| || ||||| ||||| |||||
Cp 900 GACCCATCTGAGGTATGTAGTGAGC 874

[illegible]

Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: mob.REGA+ET
High quality sequence stop: 176.

NCBI g1: 1276338

FEATURES	Location/Qualifiers
source	1..202

Organism: Homo sapiens; Vector: pT713D (Pharmacia)
/note="Organ: Liver and Spleen; Vector: pT713D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACCTGGAGCAATTAATTAAAGATCTTTTTTTTTTTTTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT713 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."

```

/clone_id="2032"
/clone_lib="Soares fetal liver spleen INFUS"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
<1...>202

```

mrna	<1..>202				
BASE COUNT	80 a	33 c	43 g	44 t	2 others
ORIGIN					
Query Match	2.58;	Score 21;	DB 147;	Length 202;	
Best Local Similarity	88.9%;	Pred. No. 1.06e-03;			
Matches	24; Conservative	0; Mismatches	3; Indels	0; Gaps	0;

```
Db      19 gaccatacctgagtgtaagtatgac 45  
        ||| | | | | | | | | |  
Cp     900 GACCCATCTGAGGTATGGTAGTGAGC 874
```

RESULT	13	T47192	301 bp	mrna	EST	08-FEB-1995
LOCUS		yb55e11.r1	Homo sapiens cDNA clone 74924 5' similar to gb:L11566			
DEFINITION		60S RIBOSOMAL PROTEIN L18 (HUMAN).				
ACCESSION		T47192				
NID		9649174				
KEYWORDS		EST.				
SOURCE		human clone-74924 library=Stratagene ovary (#937217)				

vector-Bluescript SK host-SOLR cells (kanamycin resistant)
primer-M13Rp1 Rsite1-EcoRI Rsite2-XhoI Cloned unidirectionally.
primer: Oligo dT. Total ovary tissue, normal 49 year old caucasian
female. Average insert size: 0.8 kb; Uni-ZAP XR Vector; 5' adaptor
sequence: 5'-GAATTCGCGACAG-3', 3' adaptor sequence:
5'-CTCAGATTTTTTTTTTTT-3'.

ORGANISM Homo sapiens; Chordata; Vertebrata; Gnathostomata; Mammalia
Eucaryotae; Metazoa; Eucaryotae; Metazoa; Eucaryotae; Metazoa;
Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS

1 (bases 1 to 301)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevasakis, E.,
Waterston, R., Willamson, A., Wohlmann, P. and Wilson, R.

TITLE	WASHU-MERCK EST Project
JOURNAL	Unpublished (1995)
COMMENT	Other_ESTs: yb53e11.s1

Contact: Wustl
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 294
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; con
IMAGE Consortium (info@image.llnl.gov) for further info

FEATURES	Location/Qualifiers
source	1..301

BASE COUNT	64	a	71 c	81 g	72 t	13 others
ORIGIN						

Query Match	2.5%;	Score 21;	DB 114;	Length 301;
Best Local Similarity	68.6%;	Pred. No. 1.06e-03;		
Matches	35;	Mismatches 16;	Indels 0;	Gaps 0.

Db 155 cggagatgaagcttcctcgtggtccggtgaacaagaangcgcttgttgtc 205
||| ||||||| ||| ||| ||| ||| |||
Cp 681 CGGACATGAAGCTTCTCGCCGGGAGCTAAACGAGTCCTCGATTGTG 631

RESULT	14	327 bp	MRNA	EST	22-MAY-1995
LOCUS	R41555				
DEFINITION	yf88c05.s1	Homo sapiens	CDNA clone	29736 3'	
ACCESSION	R41555				
NID	9816856				

KEYWORDS
SOURCE
EST.
human clone-29736 library-Soares infant brain 1N1B vector-Lafmld BA host-DH10B (ampicillin resistant) primer-Promega -21ml3 RsaI-e1-Not I RsaI-e2-Hind III whole brain from a 73 days post natal female. 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', AACTGGAGAAATTCGCGCCGACGAGAAATTTTCTTTTCTTTT 3']; double-stranded cDNA was ligated to Hind III adaptors (Pharmacia), digested with Not I and directionally cloned into the Not I and Hind III sites of the Lafmld BA vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo.

ORGANISM

ORGANISM
HOMO SAPIENS
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomate;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Homiidae; Homo.

REFERENCE AUTHORS	1 (bases 1 to 327) Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., to W. J. Jones G. Marra M.
----------------------	---

TITLE The Washu-Merck EST Project

JOURNAL Unpublished (1995)
COMMENT GDB: G00-402-083
Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
High quality sequence stops: 127
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
location/Qualifiers
1..327
/organism="Homo sapiens"
/clone="29736"
/note="human"
BASE COUNT 73 a 69 c 80 g 100 t 5 others
ORIGIN
Query Match 2.5%; Score 21; DB 83; Length 327;
Best Local Similarity 80.0%; Pred. No. 1.06e-03;
Matches 28; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Db 51 aaattactaaggaagacctgagcaaatgtgat 85
||||| ||||||| ||| |||||||
QY 192 AAATTTTAGAGAGAACTGAAAAATATGTGAT 226
RESULT 15
LOCUS HSC3AD011 336 bp RNA EST 21-SEP-1995
DEFINITION H. sapiens partial cDNA sequence; clone c-3ad01.
ACCESSION F12463
NID 9708455
KEYWORDS partial cDNA sequence; transcribed sequence fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondria eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 336)
AUTHORS Genexpress.
TITLE Direct Submission
JOURNAL Submitted (19-JAN-1995) Genethon, B.P. 60, 91002 Evry Cedex France
and Genetique Molculaire et Biologie du developpement, CNRS UPR420
B.P. 8, 94801 Villejuif Cedex France. E-mail:
genexpress@genethon.fr
2 (bases 1 to 336)
REFERENCE Genexpress.
AUTHORS The Genexpress cDNA program
TITLE Unpublished
JOURNAL 3 (bases 1 to 336)
REFERENCE 3 (bases 1 to 336)
AUTHORS Auffray,C., Behar,G., Bois,F., Bouchier,C., da Silva,C.,
Devignes,M.D., Duprat,S., Houlgatte,R., Jumeau,M.N., Lamy,B.,
Lorenzo,F., Mitchell,H., Mariage-Samson,R., Pietu,G., Pouliot,Y.,
Sebastiani-Kabaktchis,C. and Tessier,A.
TITLE IMAGE: molecular integration of the analysis of the human genome
and its expression
JOURNAL C. R. Acad. Sci. III, Sci. Vie 318 (2), 263-272 (1995)
MEDLINE 95277534
COMMENT Cloning_method: total mRNA was oligo-(dT) primed and directionally
cloned 5' -> 3' into the HindIII -> NotI sites of the lacmid BA
vector;
Sequencing_method: single read, full automatic;
Primer: M13_reverse
cDNA sequence colinear to mRNA
Stretch_removed: nothing
Normalization_method: Bento Soares, P.N.A.S. 91:9228-9232(1994);
Genexpress_library_idt: C;
Genexpress_sequence_idt: y1c-3ad01.
FEATURES
SOURCE Location/Qualifiers
1..336

/organism="Homo sapiens"
/isolate="muscular atrophy patient"
/clone_lib="normalized infant brain cDNA from B.Soaers,
psychiatry Dept. Columbia University USA"
/sex="female"
/tissue_type="total brain"
/dev_stage="3 months old"
BASE COUNT 82 a 78 c 103 g 68 t 5 others
ORIGIN
Query Match 2.5%; Score 21; DB 44; Length 336;
Best Local Similarity 85.7%; Pred. No. 1.06e-03;
Matches 24; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
Db 212 tncggaagatgaagcttcctgcgcggga 239
||||| ||||||| ||| |||||||
Cp 683 TGGGACATGAAGCTTCTGCCCGGA 656

Search completed: Tue Feb 11 00:04:32 1997
Job time : 457 secs.

BEST AVAILABLE COPY

CC	CLONE:	methyladenosine phosphatase
CC	FEATURE:	
CC	NAME/KEY:	CDS
CC	LOCATION:	1..2763
CC	Sequence	2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
CC	Query Match	28.3%; Score 240; DB 10; Length 2763;
CC	Best Local Similarity	99.2%; Pred. No. 9.33e-177;
CC	Matches	242; Conservative 0; Mismatches 2; Indels 0; Gaps
Db	962 AGCTTCTTAGAGACTGCTAAGAAGCTAGAGACTCCGGTCACACTCAAAGSGGACAATGG	1021
Oy	570 AGGTTCATTATAGAGACTGCTAAGAAGCTAGAGACTCCGGTGCCACTCAAAAGGGACAATGG	629
Db	1022 TCACATCGAGGAGCCTCGTTTAGCTCCCGGGCAGAAAGCTTCATGTTCCGCACCTGGG	1081
Oy	630 TCACATCGAGGAGCCTCGTTTAGCTCCCGGGCAGAAAGCTTCATGTTCCGCACCTGGG	689
Db	1082 GGGCGGATGTTATCAACATGACCACAGTTCACAGAGTGTCTTCTGCTAAGAGGCTGGAA	1141
Oy	690 GGGCGGATGTTATCAACATGACCACAGTTCACAGAGTGTCTTCTGCTAAGAGGCTGGAA	749
Db	1142 TTGTTCACGCAAGTATGCCCATGGGACACAGATTATGACTGCTGAAGGAGCAGAGGAAG	1201
Oy	750 TTGTTCACGCAAGTATGCCCATGGGACACAGATTATGACTGCTGAAGGAGCAGAGGAAG	809
Db	1202 CAGT 1205	
Oy	810 CAGT 813	
RESULT	2	
ID	US-08-176-413-1 STANDARD; DNA; UNC; 2763 BP.	
AC	xxxxxx	
DT	01-JAN-1900	
DE	Sequence 1, Application US/08176413.	
CE	Sequence 1, Application US/08176413	
CC	Patent No. 5571510	
CC	GENERAL INFORMATION:	
CC	APPLICANT: No. 5571510orl, Tsutomu	
CC	APPLICANT: Carson, Dennis A.	
CC	TITLE OF INVENTION: METHOD FOR SELECTIVE METHIONINE	
CC	TITLE OF INVENTION: STARVATION OF MALIGNANT CELLS IN MAMMALS	
CC	NUMBER OF SEQUENCES: 5	
CC	CORRESPONDENCE ADDRESS:	
CC	ADDRESSEE: Spensley Horn Jubas & Lubitz	
CC	STREET: 1880 Century Park East, Suite 500	
CC	CITY: Los Angeles	
CC	STATE: California	
CC	COUNTRY: USA	
CC	ZIP: 90067	
CC	COMPUTER READABLE FORM:	
CC	MEDIUM TYPE: Floppy disk	
CC	COMPUTER: IBM PC compatible	
CC	OPERATING SYSTEM: PC-DOS/MS-DOS	
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25	
CC	CURRENT APPLICATION DATA:	
CC	APPLICATION NUMBER: US/08/176,413	
CC	FILING DATE:	
CC	CLASSIFICATION: 424	
CC	ATTORNEY/AGENT INFORMATION:	
CC	NAME: Howells, Stacy L.	
CC	REGISTRATION NUMBER: 34,842	
CC	REFERENCE/DOCKET NUMBER: PD2864	
CC	TELECOMMUNICATION INFORMATION:	
CC	TELEPHONE: (619) 455-5100	
CC	TELEFAX: (619) 455-5110	
CC	INFORMATION FOR SEQ ID NO: 1:	
CC	SEQUENCE CHARACTERISTICS:	
CC	LENGTH: 2763 base pairs	
CC	TYPE: nucleic acid	
CC	STRANDEDNESS: single	
CC	TOPOLOGY: linear	

```
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ

Query Match      28.3%; Score 240; DB 5; Length 2763;
Best Local Similarity 99.2%; Pred. No. 9.33e-177;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db    962 AGGTCCTATAGAGACTGCTAAGAAGTAGACTCCGGTGCACCTCAAGGGCACAATGG 1021
      |||
QY    570 AGGTCTTATAGAGACTGCTAAGAAGTAGACTCCGGTGCACCTCAAGGGCACAATGG 629
      |||

Db    1022 TCACATCGAGGGACCTCGTTTAACTCCCGGCAGAAAGCTTCATGTTCCGCACCTGGG 1081
      |||
QY    630 TCACATCGAGGGACCTCGTTTAACTCCCGGCAGAAAGCTTCATGTTCCGCACCTGGG 689
      |||

Db    1082 GGGCGATGTTATCAACATGACCACAGAGTGTTCTTCTGTAAGAGGCTGGAA 1141
      |||
QY    690 GGGCGATGTTATCAACATGACCACAGAGTGTTCTTCTGTAAGAGGCTGGAA 749
      |||

Db    1142 TTGTACGCCAGATATGCCCATGGGCACAGATTATGACTGCTGAAGAGCACGAGAGAG 1201
      |||
QY    750 TTGTACGCCAGATATGCCCATGGGCACAGATTATGACTGCTGAAGAGCACGAGAGAG 809
      |||

Db    1202 CAGT 1205
      |||
QY    810 CAGT 813

RESULT          3
ID             PCT-US94-14920-1 STANDARD; DNA; UNC; 2763 BP.
AC            xxxxxx
DT           01-JAN-1900
DE     Sequence 1, Application PC/TUS9414920.
CC     Sequence 1, Application PC/TUS9414920
CC     GENERAL INFORMATION:
CC     APPLICANT: THE REGENTS OF THE UNIVERSITY
CC     APPLICANT: OF CALIFORNIA
CC     TITLE OF INVENTION: METHOD FOR DETECTION OF
CC     TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHATASE DEFICIENCY IN MAMM
CC
ALIAN
CC     TITLE OF INVENTION: CELLS
CC     NUMBER OF SEQUENCES: 5
CC     CORRESPONDENCE ADDRESS:
CC     ADDRESSEE: Robbins, Berliner & Carson
CC     STREET: 201 N. Figueroa Street, 5th Floor
CC     CITY: Los Angeles
CC     STATE: California
CC     COUNTRY: USA
CC     ZIP: 90012
CC     COMPUTER READABLE FORM:
CC     MEDIUM TYPE: Floppy disk
CC     COMPUTER: IBM PC compatible
CC     OPERATING SYSTEM: PC-DOS/MS-DOS
CC     SOFTWARE: Patent In Release #1.0, Version #1.25
CC     CURRENT APPLICATION DATA:
CC     APPLICATION NUMBER: PCT/US94/14920
CC     FILING DATE:
CC     CLASSIFICATION:
CC     ATTORNEY/AGENT INFORMATION:
CC     NAME: Berliner, Robert
CC     REGISTRATION NUMBER: 20,121
CC     REFERENCE/DOCKET NUMBER: 5555-287
CC     TELECOMMUNICATION INFORMATION:
CC     TELEPHONE: 213-977-1001
CC     TELEFAX: 213-977-1003
CC     INFORMATION FOR SEQ ID NO: 1:
CC     SEQUENCE CHARACTERISTICS:
CC     LENGTH: 2763 base pairs
```

CC	TYPE: nucleic acid
CC	STRANDEDNESS: single
CC	TOPOLOGY: linear
CC	MOLECULE TYPE: DNA (genomic)
CC	IMMEDIATE SOURCE:
CC	CLONE: methyladenosine phosphatase
CC	FEATURE:
CC	NAME/KEY: CDS
CC	LOCATION: 1..2763
CC	Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
Query Match	28.3%; Score 240; DB 10; Length 2763;
Best Local Similarity	99.2%; Pred. No. 9.33e-177;
Matches	242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db	962 AGGTTCTTATAGAGACTGCTAAGAGACTAGACTCCGGTGCACCTCAAGGGGACATGG 1021
QY	570 AGGTCTTATAGAGACTGCTAAGAGACTAGACTCCGGTGCACCTCAAGGGGACATGG 629
Db	1022 TCACAATCGAGGAGCCTCGTTTACCTCCCGGGCAGAAAGCTTCATGTTCCGCACCTGGG 1081
QY	630 TCACAATCGAGGAGCCTCGTTTACCTCCCGGGCAGAAAGCTTCATGTTCCGCACCTGGG 689
Db	1082 GGGCGGATGTATCAACATGACACAGTTCAGAGGTGTTCTTGTCTAAGGAGGCTGAA 1141
QY	690 GGGCGGATGTATCAACATGACACAGTTCAGAGGTGTTCTTGTCTAAGGAGGCTGAA 749
Db	1142 TTTGTACGCAAGTATCGCCATGGGCACAGATTATGACTGCTGGAAGAGCAGGAAG 1201
QY	750 TTTGTACGCAAGTATCGCCATGGGCACAGATTATGACTGCTGGAAGAGCAGGAAG 809
Db	1202 CAGT 1205
QY	810 CAGT 813

RESULT 4
 ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
 AC xxxxxx
 DT 01-JAN-1900
 DE Sequence 5, Application US/08238163.
 CC Sequence 5, Application US/08238163
 CC Patent No. 5569830
 CC GENERAL INFORMATION:
 CC APPLICANT: BENNETT, Alan
 CC APPLICANT: LABAVITCH, John M.
 CC APPLICANT: POWELL, Ann
 CC APPLICANT: STOTZ, Henrik
 CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
 CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL
 DISEASE
 CC NUMBER OF SEQUENCES: 24
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Townsend and Townsend Kourie and Crew
 CC STREET: Stewart Street Tower, One Market Plaza
 CC CITY: San Francisco
 CC STATE: California
 CC COUNTRY: US
 CC ZIP: 94105-1493
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/238,163
 CC FILING DATE: 03-MAY-1994
 CC CLASSIFICATION: 800
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Bastian, Kevin L.
 CC REGISTRATION NUMBER: 34,774
 CC REFERENCE/DOCKET NUMBER: 2307E-540
 CC TELECOMMUNICATION INFORMATION:

[illegible]

RESULT 5
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5, Application US/08238163.
CC Sequence 5, Application US/08238163
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL
CC DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Kourile and Crew
CC STREET: Stewart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:


```

CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 215 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: unknown
CC      MOLECULE TYPE: protein
CC      FEATURE:
CC      NAME/KEY: misc_feature
CC      LOCATION: 1..215
CC      OTHER INFORMATION: /standard_name="Deduced amino acid
CC      OTHER INFORMATION: sequence of PGIP from bean."
CC
SQ      Sequence 215 bp; 15 A; 8 C; 25 G; 26 T; 141 other;

```

Query Match	3.3%;	Score 28;	DB 5;	Length 215;
Best Local Similarity	16.9%;	Pred. No. 2.37e-05;		
Matches	21;	Conservative	49;	Mismatches 53;
			Indels 1;	Gaps 1;

D_b

49 TYRVNDSGHNKYYSSANYNGNNVGAAKTHYHTHTNSGADSKTV-TDINSAGSTSSSN 107
|::: :: : | : | : |:| : | ::: | : | :

C_p

249 TCAGATGGCTTGCCAAATGGAGTATCACAATAATTTTTCACGTCTTCCCTCTAAAATTCT 190

D_b

108 GGTGNRSGADSYSGSCKTAMTSRNRITGTANNNAVDSRNNDASVGSDKNTKKHAKNSADG 167
|| : ||::: :: : | : | : |::: :: : |:

C_p

189 GGATCATCCAGGCCCTGTTCCAACC AATTATCAATCTTCACGGCGGTGGTGGTGCCA 130

D_b

168 KVG S 171
::|:

C_p

129 GAGG 126

```

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 100, Application PC/TUS9511934.
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennile & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC Sequence 74 BP; 6 A; 6 C; 1 G; 1 T; 60 other;

```

```

Query Match          2.9%; Score 25; DB 11; Length 74;
Best Local Similarity 9.0%; Pred. No. 1.71e-03;
Matches 6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db      6 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 65
        : : : : : : : : : : : : : : : : : : : : : :
Cp      469 CCGTCAATGAAGTGCATATAATGACATATCGCCGGCTGAATCTCCTCCCTCAAGGA 410

Db      66 ACCACAA 72
        |||||
Cp      409 GCCACAA 403

```

RESULT	7
ID	PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.

DT 01-JAN-1900
DE Sequence 99, Application PC/TUS9511934.
CC Sequence 99, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA

```

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 99:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 75 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC Sequence 75 BP; 1 A; 1 C; 7 G; 5 T; 61 other;
SQ

```

Query Match	2.98;	Score 25;	DB 11;	Length 75;
Best Local Similarity	12.58;	Pred. No. 1.71e-03;		
Matches	9;	Conservative	19;	Mismatches 44;
			Indels 0;	Gaps 0;

[illegible]

```

Db      64  NNBGGTTGTGGT  75
          :||  |||||
Cp     151  CACGGCGGTGGT  140

```

RESULT	8
ID	PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.

```

AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 98, Application PC/TUS9511934.
CC      Sequence 98, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US95/11934
CC      FILING DATE: 20-SEP-1995
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mistrock, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-196-228
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 790-9090
CC      TELEFAX: (212) 869-9741/8864
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 98:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 81 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      Sequence 81 BP; 6 A; 6 C; 4 G; 5 T; 60 other;
SQ
Query Match          2.9%; Score 25; DB 11; Length 81;
Best Local Similarity 9.0%; Pred. No. 1.71e-03;
Matches 6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db      13 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 72
Cp      469 CCTGTCATGACTGATCATATATGACATATATCGCCGGCTGATCTCCTCCCTCAAGGA 410
        : : : : : : : : : : : : : : : : : : : : : :
        Matches 6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db      73 ACCACAA 79
Cp      409 GCCACAA 403
        |||||
        |||||

RESULT 9
ID      PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.
AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 97, Application PC/TUS9511934.
CC      Sequence 97, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036

```

```
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/MS-DOS  
CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: PCT-US95/11934  
CC FILING DATE: 20-SEP-1995  
CC CLASSIFICATION:  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Mistrock, S. Leslie  
CC REGISTRATION NUMBER: 18,872  
CC REFERENCE/DOCKET NUMBER: 1101-196-228  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (212) 790-9090  
CC TELEFAX: (212) 869-9741/8864  
CC TELEX: 66141 PENNIE  
CC INFORMATION FOR SEQ ID NO: 97:  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 82 base pairs  
CC TYPE: nucleic acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: DNA (genomic)  
CC Sequence 82 BP; 1 A; 2 C; 10 G; 8 T; 61 other;  
  
Query Match 2.9%; Score 25; DB 11; Length 82;  
Best Local Similarity 12.5%; Pred. No. 1.71e+03;  
Matches 9; Conservative 19; Mismatches 44; Indels 0; Gaps 0;  
  
Db 11 AGNNBNNBNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNB 70  
|| : : : : : : : : : : : : : : : :  
CP 211 AGTTCCTCCTTAATAATTGTGATCATCGAGGCCTGTTCACCAATATTCCAATCTT 152  
  
Db 71 NNBGGTTGTGT 82  
:|| ||| |  
CP 151 CACGCCGCTGTGT 140  
  
RESULT 10  
ID PCT-US95-11934-92 STANDARD; DNA; UNC; 81 BP.  
AC xxxxxx  
DT 01-JAN-1900  
DE Sequence 92, Application PC/TUS9511934.  
Sequence 92, Application PC/TUS9511934  
GENERAL INFORMATION:  
APPLICANT: Cytozen Corporation  
TITLE OF INVENTION: Antigen Binding Peptides (Abltides) From  
TITLE OF INVENTION: Peptide Libraries  
NUMBER OF SEQUENCES: 103  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/11934  
FILING DATE: 20-SEP-1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-196-228  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-9741/8864
```

```
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 92:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 81 base pairs
CC          TYPE: nucleic acid
CC          STRANDEDNESS: single
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      Sequence 81 BP; 3 A; 5 C; 6 G; 4 T; 63 other;

Query Match          2.8%; Score 24; DB 11; Length 81;
Best Local Similarity 7.5%; Pred. No. 6.82e-03;
Matches 5; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

DB      7 CTCGAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 66
OY      412 CTCGAGGAGGAGATTCAGCCCGCGCATATTGTCAATTATGATCAGTTCAATGACAGGAC 471

DB      67 BNNBNNB 73
OY      472 CACTATG 478

RESULT 11
ID      PCT-US95-11934-97 STANDARD; DNA; UNC; 82 BP.
AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 97, Application PC/TUS9511934.
CC      Sequence 97, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US95/11934
CC      FILING DATE: 20-SEP-1995
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mistrock, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-196-228
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 790-9090
CC      TELEFAX: (212) 869-9741/8864
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 97:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 82 base pairs
CC          TYPE: nucleic acid
CC          STRANDEDNESS: single
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      Sequence 82 BP; 1 A; 2 C; 10 G; 8 T; 61 other;

Query Match          2.8%; Score 24; DB 11; Length 82;
Best Local Similarity 7.5%; Pred. No. 6.82e-03;
Matches 5; Conservative 20; Mismatches 42; Indels 0; Gaps 0;

DB      7 CTCGAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 66
OY      412 CTCGAGGAGGAGATTCAGCCCGCGCATATTGTCAATTATGATCAGTTCAATGACAGGAC 471
```

```
OY      412 CTCGAGGAGGAGATTCAGCCCGCGCATATTGTCAATTATGATCAGTTCAATGACAGGAC 471
DB      67 BNNBNNB 73
OY      472 CACTATG 478

RESULT 12
ID      PCT-US95-11934-94 STANDARD; DNA; UNC; 74 BP.
AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 94, Application PC/TUS9511934.
CC      Sequence 94, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC      TITLE OF INVENTION: Peptide Libraries
CC      NUMBER OF SEQUENCES: 103
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Pennie & Edmonds
CC      STREET: 1155 Avenue of the Americas
CC      CITY: New York
CC      STATE: New York
CC      COUNTRY: USA
CC      ZIP: 10036
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: PCT/US95/11934
CC      FILING DATE: 20-SEP-1995
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Mistrock, S. Leslie
CC      REGISTRATION NUMBER: 18,872
CC      REFERENCE/DOCKET NUMBER: 1101-196-228
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (212) 790-9090
CC      TELEFAX: (212) 869-9741/8864
CC      TELEX: 66141 PENNIE
CC      INFORMATION FOR SEQ ID NO: 94:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 74 base pairs
CC          TYPE: nucleic acid
CC          STRANDEDNESS: single
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
SQ      Sequence 74 BP; 3 A; 4 C; 3 G; 1 T; 63 other;

Query Match          2.7%; Score 23; DB 11; Length 74;
Best Local Similarity 4.7%; Pred. No. 2.64e-02;
Matches 3; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

DB      3 GAGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNN 62
OY      415 GAGGAGGAGATTCAGCCCGCGCATATTGTCAATTATGATCAGTTCAATGACAGGAC 474

DB      63 BNNB 66
OY      475 TATG 478

RESULT 13
ID      PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC      xxxxxx
DT      01-JAN-1900
DE      Sequence 100, Application PC/TUS9511934.
CC      Sequence 100, Application PC/TUS9511934
CC      GENERAL INFORMATION:
CC      APPLICANT: Cytogen Corporation
CC      TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
```


Tue Feb 11 11:27:33 1997

US-08-674-311-1-02.rn

Page 8

CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 81 BP; 6 A; 6 C; 4 G; 5 T; 60 other;

Query Match	2.7%	Score 23;	DB 11;	Length 81;
Best Local Similarity	7.58;	Pred. No. 2.64e-02;		
Matches	5;	Conservative	20;	Mismatches 42;
			Indels	0;
			Gaps	0;

D6 **7** TCTAGAVNNVNNVNNTNVNNVVNNNVTNNVNNNTNVNNNVNN 66

111
128 TCTGGCACCACCACCGCCGTGAAGATTGAATAATTGGTGAACAGGCCCTGATGAT 187

Db 67 VNNVNNA 73

QY 188 CCAGAAA 194

Search completed: Tue Feb 11 00:05:28 1997
Job time : 37 secs.

DB 182 GATGATCCAGAAATTTAGAGGAGAACTGAAAAATATGTGATACTCCATTGGCAAG 241
| | | | |
QY 182 GATGATCCAGAAATTTAGAGGAGAACTGAAAAATATGTGATACTCCATTGGCAAG 241
DB 242 CCATCTGATGCTTAAATTTGGGAGATAAAAATGTGATGTCATCCTCTGCAAG 301
| | | | |
QY 242 CCATCTGATGCTTAAATTTGGGAGATAAAAATGTGATGTCATCCTCTGCAAG 301
DB 302 CATGGAAGGACACACCATCATGCTTCAAGGTCAACTACAGGCGAATCTGGCT 361
| | | | |
QY 302 CATGGAAGGACACACCATCATGCTTCAAGGTCAACTACAGGCGAATCTGGCT 361
DB 362 TTGAAGAGAGGGCTGTACACATGTCATAGTGACACAGCTTGTGGCTCTTGAGGAG 421
| | | | |
QY 362 TTGAAGAGAGGGCTGTACACATGTCATAGTGACACAGCTTGTGGCTCTTGAGGAG 421
DB 422 GAGATTGAGCCCGCGATATGTCATTTATGATCAGTTGATGACAGGACCATATGAGA 481
| | | | |
QY 422 GAGATTGAGCCCGCGATATGTCATTTATGATCAGTTGATGACAGGACCATATGAGA 481
DB 482 CCTCAGTCTTCTATGATGGAAGTCATCTTGTGCCAGAGAGTGTGCCATATTCATG 541
| | | | |
QY 482 CCTCAGTCTTCTATGATGGAAGTCATCTTGTGCCAGAGAGTGTGCCATATTCATG 541
DB 542 GCTGAGCCGTTTGGCCCAAAAGAGAGGTTCTTATAGACTGTCTAAGAGCTAGGA 601
| | | | |
QY 542 GCTGAGCCGTTTGGCCCAAAAGAGAGGTTCTTATAGACTGTCTAAGAGCTAGGA 601
DB 602 CTCCGCTGCCACTCAAGGGGACATGTCACATCGAGGACCTCGTTTAGCTCCGG 661
| | | | |
QY 602 CTCCGCTGCCACTCAAGGGGACATGTCACATCGAGGACCTCGTTTAGCTCCGG 661
DB 662 GCAGAAAGCTTCATGTTCCGACCTGGGGGGGGGATGTATCAACATGACACAGTTCCA 721
| | | | |
QY 662 GCAGAAAGCTTCATGTTCCGACCTGGGGGGGGGATGTATCAACATGACACAGTTCCA 721
DB 722 GAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTAGCAAGTATGCCATGGCGAGAT 781
| | | | |
QY 722 GAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTAGCAAGTATGCCATGGCGAGAT 781
DB 782 TATGACTGCTGGAAGGACGACGAGAGCAGTTCGTTGAGCGGGCTTAAAGCCCTG 841
| | | | |
QY 782 TATGACTGCTGGAAGGACGACGAGAGCAGTTCGTTGAGCGGGCTTAAAGCCCTG 841
DB 842 AAAGAAACGCTAATAAGCCAAAGCTTACTGCTACATACCATACCTCAGATAGGCTC 901
| | | | |
QY 842 AAAGAAACGCTAATAAGCCAAAGCTTACTGCTACATACCATACCTCAGATAGGCTC 901
DB 902 ACAGATGCTCAGAAACCCTCCATACCTGAAGAATATGGCCAGTTTCTGTTTATTA 961
| | | | |
QY 902 ACAGATGCTCAGAAACCCTCCATACCTGAAGAATATGGCCAGTTTCTGTTTATTA 961
DB 962 CCAAGACAT 970
| | | | |
QY 962 CCAAGACAT 970
| | | | |
RESULT 3
ID US-08-445-648-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08445648.
CC Sequence 14, Application US/08445648
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: IBM PC compatible
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/445,648
CC FILING DATE: 22-MAY-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/227,800
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
Query Match 28.3%; Score 240; DB 59; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2.29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
DB 962 AGGTCTTATAGAGACTGCTAAGAGCTAGAGACTCCGCTGCCACTCAAGAGGACAATGG 1021
| | | | |
QY 570 AGGTCTTATAGAGACTGCTAAGAGCTAGAGACTCCGCTGCCACTCAAGAGGACAATGG 629
DB 1022 TCACAAATCAGAGGACCTCGTTTAACTCCCGGGCAGAAAGCTTCATGTTCCGCACTGGG 1081
| | | | |
QY 630 TCACAAATCAGAGGACCTCGTTTAACTCCCGGGCAGAAAGCTTCATGTTCCGCACTGGG 689
DB 1082 GGGCGATGTTATCAACATGACACAGTTCAGAGAGTGTCTTGTCTAAGAGAGCTGGA 1141
| | | | |
QY 690 GGGCGATGTTATCAACATGACACAGTTCAGAGAGTGTCTTGTCTAAGAGAGCTGGA 749
DB 1142 TTGTACGCAAGTATGCGCATGGCCACAGATTATGACTGCTGGAAGAGCAGAGAAG 1201
| | | | |
QY 750 TTGTACGCAAGTATGCGCATGGCCACAGATTATGACTGCTGGAAGAGCAGAGAAG 809
DB 1202 CAGT 1205
| | | | |
QY 810 CAGT 813
RESULT 4
ID US-08-176-855-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/08176855.
CC Sequence 1, Application US/08176855
CC GENERAL INFORMATION:
CC APPLICANT: Nobori, Tsutomu
CC APPLICANT: Carson, Dennis A.

CC APPLICANT: Takabayashi, Kenji
CC TITLE OF INVENTION: METHOD FOR DETECTION OF
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHATASE DEFICIENCY IN MAMM
ALIAN
CC TITLE OF INVENTION: CELLS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Spensley Horn Jubas & Lubitz
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/176,855
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Howells, Stacy L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD3057
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 455-5100
CC TELEFAX: (619) 455-5110
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC SEQUENCE 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 28.3%; Score 240; DB 32; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2.29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 962 AGTTCCTTATAGAGACTGCTAGAGAGGACTCCGGTCCACTCAAGGGGACATGG 1021
QY 570 AGTTCCTTATAGAGACTGCTAGAGAGGACTCCGGTCCACTCAAGGGGACATGG 629
Db 1022 TCACAATCGAGGAGCCTGTTTACGTCCTCCGGGACAGAAAGCTTCATGTTCCGACCTGGG 1081
QY 630 TCACAATCGAGGAGCCTGTTTACGTCCTCCGGGACAGAAAGCTTCATGTTCCGACCTGGG 689
Db 1082 GGGCGATGTTATCAACATGACACAGTTCAGAGGTGCTTCTGCTAAGGAGGCTGGA 1141
QY 690 GGGCGATGTTATCAACATGACACAGTTCAGAGGTGCTTCTGCTAAGGAGGCTGGA 749
Db 1142 TTTGTACGCAAGTATCGCATGGGACAGATTATGACTGCTGGAAGGAGCAGAGGAAG 1201
QY 750 TTTGTACGCAAGTATCGCATGGGACAGATTATGACTGCTGGAAGGAGCAGAGGAAG 809
Db 1202 CAGT 1205
QY 810 CAGT 813

RESULT 5
ID US-08-443-880-14 STANDARD: DNA; UNC: 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08443880.

CC Sequence 14, Application US/08443880
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/443,880
CC FILING DATE: 22-MAY-1995
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/227,800
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC SEQUENCE 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 28.3%; Score 240; DB 59; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2.29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Db 962 AGTTCCTTATAGAGACTGCTAGAGAGGACTCCGGTCCACTCAAGGGGACATGG 1021
QY 570 AGTTCCTTATAGAGACTGCTAGAGAGGACTCCGGTCCACTCAAGGGGACATGG 629
Db 1022 TCACAATCGAGGAGCCTGTTTACGTCCTCCGGGACAGAAAGCTTCATGTTCCGACCTGGG 1081
QY 630 TCACAATCGAGGAGCCTGTTTACGTCCTCCGGGACAGAAAGCTTCATGTTCCGACCTGGG 689
Db 1082 GGGCGATGTTATCAACATGACACAGTTCAGAGGTGCTTCTGCTAAGGAGGCTGGA 1141
QY 690 GGGCGATGTTATCAACATGACACAGTTCAGAGGTGCTTCTGCTAAGGAGGCTGGA 749
Db 1142 TTTGTACGCAAGTATCGCATGGGACAGATTATGACTGCTGGAAGGAGCAGAGGAAG 1201
QY 750 TTTGTACGCAAGTATCGCATGGGACAGATTATGACTGCTGGAAGGAGCAGAGGAAG 809
Db 1202 CAGT 1205
QY 810 CAGT 813

RESULT 6
ID US-08-227-800A-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800A.
CC Sequence 14, Application US/08227800A
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800A
CC FILING DATE: 14-APR-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;

Query Match 28.3%; Score 240; DB 37; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2.29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 AGGTTCTATAGAGACTGCTAGAGACTCCGGTCCACTCAAGGGGACATGG 1021
|||||
QY 570 AGGTTCTATAGAGACTGCTAGAGACTCCGGTCCACTCAAGGGGACATGG 629
|||||
Db 1022 TCACATCGAGGAGCTCGTTTAGCTCCCGGCGAGAAAGCTTCATGTTCCGCACCTGGG 1081
|||||
QY 630 TCACATCGAGGAGCTCGTTTAGCTCCCGGCGAGAAAGCTTCATGTTCCGCACCTGGG 689
|||||
Db 1082 GGGCGGATGTATCAACATGACACAGTTCAGAGGTTCTTGTCTAAGAGGCTGGAA 1141
|||||
QY 690 GGGCGGATGTATCAACATGACACAGTTCAGAGGTTCTTGTCTAAGAGGCTGGAA 749
|||||
Db 1142 TTGTTACGCAAGTATCGCCATGGGCGACAGATTATGACTGCTGGAAGGACGAGGAAG 1201
|||||
QY 750 TTGTTACGCAAGTATCGCCATGGGCGACAGATTATGACTGCTGGAAGGACGAGGAAG 809
Db 1202 CAGT 1205

QY 810 CAGT 813
|||||
RESULT 7
ID US-08-227-800-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800.
CC Sequence 14, Application US/08227800
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: METHODS FOR TREATMENT AND DETECTION OF
CC TITLE OF INVENTION: SUSCEPTIBILITY TO, AND PROGRESSION OF, CANCER
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: US
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD3459
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/455-5100
CC TELEFAX: 619/455-5110
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;

Query Match 28.3%; Score 240; DB 37; Length 2763;
Best Local Similarity 99.2%; Pred. No. 2.29e-222;
Matches 242; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 962 AGGTTCTATAGAGACTGCTAGAGACTCCGGTCCACTCAAGGGGACATGG 1021
|||||
QY 570 AGGTTCTATAGAGACTGCTAGAGACTCCGGTCCACTCAAGGGGACATGG 629
|||||
Db 1022 TCACATCGAGGAGCTCGTTTAGCTCCCGGCGAGAAAGCTTCATGTTCCGCACCTGGG 1081
|||||
QY 630 TCACATCGAGGAGCTCGTTTAGCTCCCGGCGAGAAAGCTTCATGTTCCGCACCTGGG 689
|||||
Db 1082 GGGCGGATGTATCAACATGACACAGTTCAGAGGTTCTTGTCTAAGAGGCTGGAA 1141
|||||
QY 690 GGGCGGATGTATCAACATGACACAGTTCAGAGGTTCTTGTCTAAGAGGCTGGAA 749
|||||
Db 1142 TTGTTACGCAAGTATCGCCATGGGCGACAGATTATGACTGCTGGAAGGACGAGGAAG 1201
|||||
QY 750 TTGTTACGCAAGTATCGCCATGGGCGACAGATTATGACTGCTGGAAGGACGAGGAAG 809

Db 1202 CAGT 1205
OY 810 CAGT 813

RESULT 8
ID US-08-220-691-5097 STANDARD; DNA; UNC; 495 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5097, Application US/08220691.
CC Sequence 5097, Application US/08220691
CC GENERAL INFORMATION:
CC APPLICANT: Haseltine, William, et. al.
CC TITLE OF INVENTION: Human Genes, Sequences, and
CC NUMBER OF SEQUENCES: 12049
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
CC STREET: 6 Becker Farm Road
CC CITY: Roseland
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
CC COMPUTER: HP Vectra 486/33
CC OPERATING SYSTEM: MSDOS version 5.0
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/220,691
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Olstein, Elliot
CC REGISTRATION NUMBER: 24,025
CC REFERENCE/DOCKET NUMBER: 325800-71
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 994-1700
CC TELEFAX: (201) 994-1744
CC INFORMATION FOR SEQ ID NO: 5097:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC Sequence 495 BP; 97 A; 96 C; 127 G; 131 T; 44 other;

Query Match 20.4%; Score 173; DB 37; Length 495;
Best Local Similarity 81.3%; Pred. 4.96e-150;
Matches 260; Conservative 0; Mismatches 50; Indels 10; Gaps 10;
Db 109 ATGGCTCTGGCACCACACCGCGTGTAGATTGNAATATTGGTGAACAGGCC 168
OY 122 ATGGCTCTGGCACCACACCGCGGTG-AGATTGG-AATTAATTGGTGAACAGGCC 179
Db 169 TGGATGATCCAGAA-TTTAGAGAAGAACTGCCAATATGTGATCTCCATTGGCA 227
OY 180 TGGATGATCCAGAAATTTAGAGAAGAACTGAAATATGTGATCTCCATTGGCA 239
Db 228 AGCCATCTGATGCTTAATTTGGGGGAAGNTAAATGTTGATTGCGTNCCTTGGC 287
OY 240 AGCCATCTGATGCTTAATTTGGGG-AGATATAAAATGTTGATGCACTCCTTG-C 297
Db 288 AAAGCATGTTAGNAGNACACCATCANGGCTTTAAAGGTTCAATTACCAGGGGANNNT 347
OY 298 AA-GGCATGGAAGCAGACACCATCATG-CCTTCAAAGGT-CAACTACAGGCGACAT 354
Db 348 TGGGNTTNNNAAGAA-AGGTTTNNNAACCTTTTNGNGACCCNANNTTNNNGTNNNT 406

OY 355 CTGGCTTTGAAGAGAGGCGTGTACACATGTATAGTACACACAGCTT-GTGGCTCCT 413
Db 407 TGGGGGGGNTTTNACC 426
OY 414 TGAGGAGAGATTACGCC 433

RESULT 9
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,190
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

Query Match 11.7%; Score 99; DB 66; Length 578;
Best Local Similarity 12.4%; Pred. No. 3.37e-72;
Matches 70; Conservative 266; Mismatches 222; Indels 7; Gaps 7;
Db 3 SARVYTS-SRDSGRVYGANMVSSTSSDKVKAHGSRSVTYVKAAGKVDVSGKKSK 61
OY 349 GAACATCTGGGCTTTGAAGAGAGGCGTGTACACATGTATAGTACACACAGCTTG 408
Db 62 DGS DKYSHKSKSKNVVTKKRYVADSKKRGCMKXKSHDMDGTWNNNAKTSWSKGGT 121
OY 409 CTCCTTGAGGAGAGATCAGCCCGCGAT-ATGTCAATATATGATCAAGTTCATGACA 467
Db 122 RDSADYHDYVGTDSKWKHSTSVTKVAHTWNRVYVAKHNRSCHTDNVKGTANTGNKAVG 181
OY 468 GGACACATATGAGACCTCAGCTCTTATGATGAGAGTCAATCTTGTGCGCAGAGAGTGT 527
Db 182 VSMNGTSVNSHTSGSKRRNNYMRAGDKSNTHRHWGDNRYVDTWATKKYADSHDR 241
OY 528 GCCATATTCATATGCTGAGCCGTTTGGCCCCCAAAAGAGAGAGGTTCTTATAGAGACTG 587
Db 242 RKVHTATYRTRDKYATTKATGMKYNWCDRWKSYVHVCSYGSISDMTSDHSV-ATA 300

QY 588 CTAAGAA-GCTAGACTCCGGTCC-CACCTCAAGGGGACATGTCACAAATCGAGGACC 645
Db 301 GVTSVSKNGTVDGRCYATKTKSTKYHSSCSVKSNGSGVVKGTKKSDYDHSKSSDS 360
QY 646 TCGTTTGTAGCTCCCGGCGAAGAACTTCATGTTCCGACCTGGGGCGGATGTATCAA 705
Db 361 YGCGARATTTTTHGTHGKTSGKTRKYDVKTRDSSGK-TKSTSHDKMNTSRACSSST 419
QY 706 CATGACCACAGTTCAGAGGTGCTTCTGCTAAGAGAGCTGGAATTTGTTACCAAGTAT 765
Db 420 NNYMGVGMHVKTSDTAWSYDKDSGCRGSTGSKKSTANRGRTSRSDGNAGDITKMYG 479
QY 766 CGCCATGGCGACAGATTATGACTGCTGGAAGAGCAGCAGAGAGAGAGAGAGAGAGAG 825
Db 480 SSSKARKDSKMRKCGSSVTKAADRGGKVARRSTGSSSAGRAAGGDKSGKTVSSAV-AK 538
QY 826 GGTCTTAAGACCCCTGAAGAAACGCTAATAAGCCAAAGCTTACTGCTCATAACCAT 885
Db 539 RKRSNTTTRVKSVAHHSKRDYRD 563
QY 886 ACCTCAGATAGGCTCAGAGATG 910

RESULT 10
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC TITLE OF INVENTION: 5-PHOSPHATASE
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,190
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SO Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

Query Match 8.7%; Score 74; DB 66; Length 578;
Best Local Similarity 11.7%; Pred. No. 5,66e-47;
Matches 48; Conservative 207; Mismatches 147; Indels 9; Gaps 9;

Db 3 SARVTSRSDSGRRVGNMNVSKSTSSSDKYKAHSHRSVTYKASGKMKVDVSGKKKSKD 62
Cp 679 GAACATGAAGCTTCTCCCGGAGAGCTAAACGAGGCTCCCTCGATTG-TGACCATTGTCC 621
Db 63 GSDKSHKKSKKNKYTKRKYVADSKRGCN-KHSDMTGTWNGNAKKTSMKSGKT 121
Cp 620 CCTTTGAGTGGCACCAGGAGCTCTAGCTTCTTACAGAGTCTCTATA-AGAACCCTCTCGTT 562
Db 122 RDDSADYHDYVGTDSKMKHSTSVTKVAHTVNRVYAKHNRSCHTDNVKTGANTGNKGA 181
Cp 561 TTGGGGCAAAACGGCTC-AGCCATTGAATATGGCACACTCTCTGGCACAAAGATGACT 503
Db 182 VSMNGTSGVNSHTSGSKRRNNMNRAGDKSNTHTHWGDNRYVDTWATKKYADSHDTR 241
Cp 502 TCCA-TCATAGAAGAGAGTCTCATAGTGTCTCTCATATGAAGTATGAT-CAATAATG 445
Db 242 RKVHTATYRTTRDKYATKKATGMKYNMCDRVKSYHVCSYGSNTSDMTSDHVSATAG 301
Cp 444 ACATATCGCCGGGCTGATCTCTCTCCCTCAAGAGCAGCAGAGAGAGAGAGAGAGAG 386
Db 302 VTSVSKNGTVDGRCYATKTKSTKYHSSCSVKSNGSGVVKGTKKSDYDHSKSSDSY 361
Cp 385 ATGTGTACAGCCCTCTCTCTCAAGAGCCAGAGTGT-CGCTGTGATGTGACCTTGAAG 327
Db 362 GCGARATTTTTHGTHGKTSGKTRKYDVKTRDSSGKTKSTSHDKMNTSR 412
Cp 326 GCATGATGTGTGCTGCTTCATGCTTCAAGAGGAG-GATGCAATCAACA 277

RESULT 11
ID US-08-446-939-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446939.
CC Sequence 8, Application US/08446939
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Milltia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,939
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FX
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240

Tue Feb 11 11:27:35 1997

US-08-674-311-1-02.rnp

Page 8

```

CC      TELEFAX: 617-861-9540
CC      INFORMATION FOR SEQ ID NO: 8:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 383 base pairs
CC          TYPE: nucleic acid
CC          STRANDEDNESS: double
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;
SQ
  Query Match      7.8%; Score 66; DB 59; Length 383;
  Best Local Similarity 11.9%; Pred. No. 4,01e-39;
  Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
  1 MTSSGVCCHSGAAGNHWYTRASTHYKSYNNDMTVDNNGAVKCKDVRSTCDNKSQMSNCS 60
  697 ATCCGCCCCCGCAGCTGCGGACATGAAGCTTTCTGCCCGGAGCTAAACGAGGTCCTC 638
  61 TSCGCAVAVWRKNDNTVCHDKYHDDAASKCMKKKGTCMSSCDNDNSYNTSNDVVTG 120
  637 GATGTGACCATTTGTCCTTGTGAGTGGCACCGGAGCTCTAGCTTTAGCAGTCTC-TA 579
  121 SGVASYCYRVNRKSSWTGKTRKMSHCADDRSDSSITCAANNHNTDVGKRAVYKAKN 180
  578 T-AAGAACCCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATTGGAATA-TGGCACACTCC 521
  181 TSTVAVKYDYASWKDKDSDNKHNTARKTKGYWTAHAKGNYTRHVSNDRVNGSSARGSH 240
  520 TCTGGACACAAGATGACTTCATCATGAGAAGACTGAGGCTCATAGTGTCTGCTCAAT 461
  241 HSDHTGCRKVVHRDKSSNVKNDTCCDCDSRGYSVDDANSVGTARYMAVSRMNNASKTD 300
  460 GAACGTGATCAATAATGACAAATATCGCGGGCTGAATCTTCCTCCT-CAAGAGCCACAAG 402
  301 VYSMAVMTSRCNNAVGYKDYGSKVRDYSMKDNVRDRGTRNSSMNHGAVCTTCWDHDART 360
  401 CTGTGTACTATGACATGTGTACAGCCCTCTTCCTTCAAGCCACAGATGTTCGCTGCT 342
  361 A 361
  341 A 341
  RESULT 12
  ID US-08-311-703-8 STANDARD; DNA; UNC; 383 BP.
  AC xxxxxx
  DT 01-JAN-1900
  DE Sequence 8, Application US/08311703.
  CC Sequence 8, Application US/08311703
  CC GENERAL INFORMATION:
  CC APPLICANT: Lin, Herbert Y.
  CC APPLICANT: Wang, Xiao-Fan
  CC APPLICANT: Weinberg, Robert A.
  CC APPLICANT: Lodish, Harvey F.
  CC TITLE OF INVENTION: TGF-Beta Type Receptor cDNAs Encoded
  CC TITLE OF INVENTION: Products and Uses Therefor
  CC NUMBER OF SEQUENCES: 8
  CC CORRESPONDENCE ADDRESS:
  CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
  CC STREET: Two Millita Drive
  CC CITY: Lexington
  CC STATE: Massachusetts
  CC COUNTRY: U.S.A.
  CC ZIP: 02173
  CC COMPUTER READABLE FORM:
  CC MEDIUM TYPE: Floppy disk
  CC COMPUTER: IBM PC compatible
  CC OPERATING SYSTEM: PC-DOS/MS-DOS
  CC SOFTWARE: PatentIn Release #1.0, Version #1.25
  CC CURRENT APPLICATION DATA:
  CC APPLICATION NUMBER: US/08/311,703
  CC FILING DATE:
  CC CLASSIFICATION: 530

```

```
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US 07/786,063
CC      FILING DATE: 31-OCT-1991
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Granahan, Patricia
CC      REGISTRATION NUMBER: 32,227
CC      REFERENCE/DOCKET NUMBER: WHI91-09
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 617-861-6240
CC      TELEFAX: 617-861-9540
CC      INFORMATION FOR SEQ ID NO: 8:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 383 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: double
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: DNA (genomic)
CC      Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;
SQ

Query Match          7.8%; Score 66; DB 46; Length 383;
Best Local Similarity 11.9%; Pred. No. 4.0le-39;
Matches    43; Conservative 165; Mismatches 149; Indels   4; Gaps   4;
```

CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,936
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FY
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 7.8%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGCVCHGSAAGWHVWTRASTHVSVNNDVTDNNGAVKCKDVRSTCDNKSMSNCS 60
CP 697 ATCCGCCCCCAGGTGCGAACAATGAGCTTTCTGCCCCGGAGCTAAACGAGGTCCCTC 638
Db 61 TSCKVCAVVRKNDNTTVCHDKYHDDAASKCMKKKGTMCSSSDCNDNSYNTSNDVVTG 120
CP 637 GATTGTGACCAATTGTCCCTTTGAGTGGCACCAGAGTCTAGCTTCTTAGCAGTCTC-TA 579
Db 121 SGVASVYCYRVNRKSSWTGTRKMSHCADDRSDSSTCANNNHNTDTVGKRAYKAKKN 180
CP 578 T-AAGAACCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATTGGAATA-TGGCACACTCC 521
Db 181 TSTVAVKYDHYASWKDRKSDNKHNTARKTKYWTAAHAKGNYTRHVSWDRNVGSSARGSH 240
CP 520 TCTGGCACAGAATGACTTCCATCATAGAGAGCTGAGGTCTCATAGTGTCTCTCAAT 461
Db 241 HSDHTGGRKMYHRDKSSNVKNDTCCDGSRGYSSVDANSVGVTARYMAVSRMNNASKTD 300
CP 460 GAACGTGATCAATATGACAATATCGCCGGCTGAATCTCTCCCT-CAAGAGCCACAAG 402
Db 301 VYSMAVWMTSRCAVGVKDYGSKVRDVYSMKDNVDRGRNNSWNHGMVCTTCWDHART 360
CP 401 CTGTGTACATATGACATATGTATACAGCCCTCTTCTTCAAGGCCAGATGTTCGCCTGCT 342
Db 361 A 361
CP 341 A 341

RESULT 14
ID US-08-446-937-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446937.
CC Sequence 8, Application US/08446937
CC GENERAL INFORMATION:

CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Miltlla Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,937
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FW
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 7.8%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 4.01e-39;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGCVCHGSAAGWHVWTRASTHVSVNNDVTDNNGAVKCKDVRSTCDNKSMSNCS 60
CP 697 ATCCGCCCCCAGGTGCGAACAATGAGCTTTCTGCCCCGGAGCTAAACGAGGTCCCTC 638
Db 61 TSCKVCAVVRKNDNTTVCHDKYHDDAASKCMKKKGTMCSSSDCNDNSYNTSNDVVTG 120
CP 637 GATTGTGACCAATTGTCCCTTTGAGTGGCACCAGAGTCTAGCTTCTTAGCAGTCTC-TA 579
Db 121 SGVASVYCYRVNRKSSWTGTRKMSHCADDRSDSSTCANNNHNTDTVGKRAYKAKKN 180
CP 578 T-AAGAACCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATTGGAATA-TGGCACACTCC 521
Db 181 TSTVAVKYDHYASWKDRKSDNKHNTARKTKYWTAAHAKGNYTRHVSWDRNVGSSARGSH 240
CP 520 TCTGGCACAGAATGACTTCCATCATAGAGAGCTGAGGTCTCATAGTGTCTCTCAAT 461
Db 241 HSDHTGGRKMYHRDKSSNVKNDTCCDGSRGYSSVDANSVGVTARYMAVSRMNNASKTD 300
CP 460 GAACGTGATCAATATGACAATATCGCCGGCTGAATCTCTCCCT-CAAGAGCCACAAG 402
Db 301 VYSMAVWMTSRCAVGVKDYGSKVRDVYSMKDNVDRGRNNSWNHGMVCTTCWDHART 360

CP 401 CTGTGTCATATGACATGTGTACAGCCCTCTTCTTCAAGCCAGATGTTCCCTGGT 342
DB 361 A 361
CP 341 A 341

RESULT 15

ID US-08-451-946-8 STANDARD; DNA; UNC; 383 BP.

AC xxxxxx

DT 01-JAN-1900

DE Sequence 8, Application US/08451946.

CC Sequence 8, Application US/08451946

CC GENERAL INFORMATION:

CC APPLICANT: Lin, Herbert Y.

CC APPLICANT: Wang, Xiao-Fan

CC APPLICANT: Weinberg, Robert A.

CC APPLICANT: Lodish, Harvey F.

CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded

CC TITLE OF INVENTION: Products and Uses Therefor

CC NUMBER OF SEQUENCES: 8

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

CC STREET: Two Militia Drive

CC CITY: Lexington

CC STATE: Massachusetts

CC COUNTRY: U.S.A.

CC ZIP: 02173

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patentin Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/451,946

CC FILING DATE:

CC CLASSIFICATION: 530

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/311,703

CC FILING DATE: 23-SEP-1994

CC CLASSIFICATION: 530

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 07/786,063

CC FILING DATE: 31-OCT-1991

CC CLASSIFICATION: 530

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Granahan, Patricia

CC REGISTRATION NUMBER: 32,227

CC REFERENCE/DOCKET NUMBER: WHI91-09FV

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 617-861-6240

CC TELEFAX: 617-861-9540

CC INFORMATION FOR SEQ ID NO: 8:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 383 base pairs

CC TYPE: nucleic acid

CC STRANDEDNESS: double

CC TOPOLOGY: linear

CC MOLECULE TYPE: DNA (genomic)

CC Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 7.8%; Score 66; DB 60; Length 383;

Best Local Similarity 11.9%; Pred. No. 4.01e-39;

Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;

DB 1 MTSSGVCHSGAGNHWTRASTHVKSNNMDVTDNNGAVKCKDVRSTCDNKSMSNCS 60
CP 697 ATCCGCCCCCGAGTGGCGGAACATGAGCTTCTGCCCGGAGCTAAACGAGTCCCTC 638
DB 61 TSCKVCVAVWRKNDNTVCHDKYHDDAASKMKKKKGTMCSSSDCNDNSYNTSNDVVTG 120
CP 637 GATTGTGACCATTTGCCCTTGTAGTGACACCGGAGTCTAGCTTCTTAGCAGTCTC-TA 579

DB 121 SGVASVYCYRVNRKSSWTGKTRKMSHCADDRSDSSITCANNHNTDVTGKRAVYKAKN 180
CP 578 T-AAGAACCTCTCTGCTTTTGGGGCAAAAGGCTCAGCCATTGGAATA-TGGCACACTCC 521
DB 181 TSTVAVKYDHYASWKRKSDNKHNTARKTGKWTAAKGNVTRHVSMDRNVGSARGSH 240
CP 520 TCTGGACAAGAATGACTTCCATCATAGAAGACTGAGCTCTCATAGTGTCTGCTCAAT 461
DB 241 HSDHTGCRKMYHRDKSNVKNKDTCCCGDSRGYSYVDANSVGVTARYMVS RMNNASKTD 300
CP 460 GAACGTGATCAATAATGACAATATCGCCGGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
DB 301 VYSMAVWMTSRCAVGVKDYGSKVRDVYSMKDNVRDGRTRNSSWNHGVCTTCWDHART 360
CP 401 CTGTGTCATATGACATGTGTACAGCCCTCTTCTTCAAGCCAGATGTTCCCTGGT 342
DB 361 A 361
CP 341 A 341

Search completed: Tue Feb 11 00:26:10 1997
Job time : 1221 secs.

MIRAGE

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.
Distribution rights by IntelliGenetics, Inc.

Mpsrch_nm n.a. - n.a. database search, using Smith-Waterman algorithm

Run on: Mon Feb 10 23:14:27 1997; MasPar time 85.57 Seconds

Tabular output not generated. 733.430 Million cell updates/sec

Title: >US-08-674-311-1 *Antic 100*
Description: (1-2269) from US08674311.seq (1 of 2)
Perfect Score: 2269
N.A. Sequence: 1 GAATTCGCTCCGACACTGCT.....AAAAAAGCGGAAATTC 2269
Comp: CTTAAGCGGAGCGCTGACGA.....TTTTTTTTTCGCCCTAAG

Scoring table: TABLE default
Gap 6

Mmatch STD : Dbase 0; Query 0

Searched: 50278 seqs, 13829960 bases x 2

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:

n-issued
1:51 2:52 3:53 4:54 5:55 6:PCT90 7:PCT91 8:PCT92 9:PCT93
10:PCT94 11:PCT95 12:PCT96

Statistics: Mean 9.387; Variance 5.460; scale 1.719

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	Pred. No.
1	466	20.5	2763	10	PCT-US94-1 Sequence 1, Applicatio	0.00e+00
2	466	20.5	2763	5	US-08-176- Sequence 1, Applicatio	0.00e+00
3	466	20.5	2763	10	PCT-US94-1 Sequence 1, Applicatio	0.00e+00
4	33	1.5	215	5	US-08-238- Sequence 5, Applicatio	7.16e-05
5	31	1.4	215	5	US-08-238- Sequence 5, Applicatio	7.16e-05
6	24	1.1	74	11	PCT-US95-1 Sequence 100, Applicat	2.80e-01
7	25	1.1	74	11	PCT-US95-1 Sequence 100, Applicat	9.13e-02
8	25	1.1	75	11	PCT-US95-1 Sequence 99, Applicati	9.13e-02
9	24	1.1	81	11	PCT-US95-1 Sequence 98, Applicati	2.80e-01
10	24	1.1	81	11	PCT-US95-1 Sequence 98, Applicati	2.80e-01
11	25	1.1	81	11	PCT-US95-1 Sequence 98, Applicati	2.80e-01
12	24	1.1	82	11	PCT-US95-1 Sequence 97, Applicati	2.80e-01
13	25	1.1	82	11	PCT-US95-1 Sequence 97, Applicati	2.80e-01
14	24	1.1	82	11	PCT-US95-1 Sequence 97, Applicati	2.80e-01
15	22	1.0	3214	5	US-08-484- Sequence 17, Applicati	9.13e-02
16	23	1.0	26	12	PCT-US96-0 Sequence 269, Applicat	2.45e+00
17	22	1.0	74	11	PCT-US95-1 Sequence 94, Applicati	8.40e-01
18	23	1.0	74	11	PCT-US95-1 Sequence 94, Applicati	8.40e-01
19	22	1.0	75	11	PCT-US95-1 Sequence 94, Applicati	8.40e-01
20	22	1.0	81	11	PCT-US95-1 Sequence 92, Applicati	2.45e+00

ALIGNMENTS

RESULT	ID	Sequence	Score	Length	ID	Description	Pred. No.
20	22	1.0	397	8	PCT-US92-1	Sequence 8, Applicatio	2.45e+00
21	22	1.0	1064	10	PCT-US94-1	Sequence 15, Applicati	2.45e+00
22	22	1.0	1270	10	PCT-US94-1	Sequence 23, Applicati	2.45e+00
23	22	1.0	1841	11	PCT-US95-0	Sequence 1, Applicatio	2.45e+00
24	22	1.0	1952	5	US-08-333-	Sequence 1, Applicatio	2.45e+00
25	22	1.0	2148	8	PCT-US92-0	Sequence 12, Applicati	2.45e+00
26	23	1.0	2504	5	US-08-484-	Sequence 15, Applicati	8.40e-01
27	22	1.0	2653	5	US-08-325-	Sequence 1, Applicatio	2.45e+00
28	23	1.0	2913	5	US-08-243-	Sequence 7, Applicatio	8.40e-01
29	23	1.0	2923	5	US-08-243-	Sequence 6, Applicatio	8.40e-01
30	22	1.0	3035	10	PCT-US94-0	Sequence 1, Applicatio	2.45e+00
31	22	1.0	3074	9	PCT-US93-1	Sequence 3, Applicatio	2.45e+00
32	23	1.0	3183	5	US-08-243-	Sequence 8, Applicatio	8.40e-01
33	23	1.0	3187	11	PCT-US95-0	Sequence 5, Applicatio	8.40e-01
34	22	1.0	3333	11	PCT-US95-0	Sequence 1, Applicatio	2.45e+00
35	22	1.0	3933	5	US-08-199-	Sequence 1, Applicatio	2.45e+00
36	23	1.0	8342	10	PCT-US94-0	Sequence 63, Applicati	8.40e-01
37	23	1.0	8392	9	PCT-US93-0	Sequence 6, Applicatio	8.40e-01
38	23	1.0	8392	4	US-08-080-	Sequence 6, Applicatio	8.40e-01
39	22	1.0	19557	8	PCT-US92-0	Sequence 1, Applicatio	2.45e+00
40	21	0.9	105	4	US-07-865-	Sequence 13, Applicati	6.97e+00
41	21	0.9	2258	3	US-07-720-	Sequence 1, Applicatio	6.97e+00
42	21	0.9	2504	5	US-08-484-	Sequence 15, Applicati	6.97e+00
43	21	0.9	2511	5	US-08-363-	Sequence 1, Applicatio	6.97e+00
44	21	0.9	3138	3	US-07-867-	Sequence 4, Applicatio	6.97e+00
45	21	0.9	3291	10	PCT-US94-0	Sequence 1, Applicatio	6.97e+00

RESULT 1
ID PCT-US94-14919-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application PC/TUS9414919.
CC Sequence 1, Application PC/TUS9414919.
CC GENERAL INFORMATION:
CC APPLICANT: THE REGENTS OF THE UNIVERSITY
CC APPLICANT: OF CALIFORNIA
CC TITLE OF INVENTION: METHOD FOR SELECTIVE METHIONINE
CC NUMBER OF INVENTION: STARVATION OF MALIGNANT CELLS IN MAMMALS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Robbins, Berliner & Carson
CC STREET: 201 N. Figueroa Street, 5th Floor
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90012
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/14919
CC FILING DATE:
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Berliner, Robert
CC REGISTRATION NUMBER: 20,121
CC REFERENCE/DOCKET NUMBER: 5555-286
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 213-977-1003
CC TELEFAX: 213-977-1001
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:

CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 20.5%; Score 466; DB 10; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
|||||
QY 933 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 992
Db 2330 AGAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGGAAAT 2389
|||||
QY 993 AGAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGGAAAT 1052
|||||
Db 2390 ATGGGAAGACATGCAGCTTTCATGCCCTTGCTATCAAGAGATGTTGTAGAGAAAGAC 2449
|||||
QY 1053 ATGGGAAGACATGCAGCTTTCATGCCCTTGCTATCAAGAGATGTTGTAGAGAAAGAC 1112
|||||
Db 2450 AAGACATGTGTGTAT-AGAGACTCCTCAATGATTTAGACAACCTTCAAAATACAGAGAA 2508
|||||
QY 1113 AAGACATGTGTGTATTAGAGACTCCTGAATGATTTAGACAACCTTCAAAATACAGAGAA 1172
|||||
Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATATACATTTTAAGGGGAAAAAATAA 2565
|||||
QY 1173 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATATACATTTTAAGGGGAAAAAATAA 1232
|||||
Db 2566 CCCACCATTCCTCTCCCTTATTAATTTGCAACAATAAAGGGTGGAGGGTAACTTC 2625
|||||
QY 1233 CCCACCATTCCTCTCCCTTATTAATTTGCAACAATAAAGGGTGGAGGGTAACTTC 1292
|||||
Db 2626 TACTTTCCTATACCTGCCAAAGAAATGAGGAAGAAATGGACTCTTTGCTTATTTATGA 2685
|||||
QY 1293 TACTTTCCTATACCTGCCAAAGAAATGAGGAAGAAATGGACTCTTTGCTTATTTATGA 1352
|||||
Db 2686 TGGCACTGTAATTTGTACAGTATTTCTGGAGGGCAATTTGCTAAATGCATCAAAAGAC 2745
|||||
QY 1353 TGGCACTGTAATTTGTACAGTATTTCTGGAGGGCAATTTGCTAAATGCATCAAAAGAC 1412
|||||
Db 2746 TTAATAATACGACGT 2761
|||||
QY 1413 TTAATAATACGACGT 1428
RESULT 2
ID US-08-176-413-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/08176413.
CC Sequence 1, Application US/08176413
CC Patent No. 5571510
CC GENERAL INFORMATION:
CC APPLICANT: No. 5571510or1, Tsutomu
CC APPLICANT: Carlson, Dennis A.
CC TITLE OF INVENTION: METHOD FOR SELECTIVE METHIONINE
CC TITLE OF INVENTION: STARVATION OF MALIGNANT CELLS IN MAMMALS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Spensley Horn Juhas & Lubitz
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/176,413
CC FILING DATE:
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Howells, Stacy L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD2864
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 455-5100
CC TELEFAX: (619) 455-5110
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 20.5%; Score 466; DB 5; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
|||||
QY 933 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 992
|||||
Db 2330 AGAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGGAAAT 2389
|||||
QY 993 AGAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGGAAAT 1052
|||||
Db 2390 ATGGGAAGACATGCAGCTTTCATGCCCTTGCTATCAAGAGATGTTGTAGAGAAAGAC 2449
|||||
QY 1053 ATGGGAAGACATGCAGCTTTCATGCCCTTGCTATCAAGAGATGTTGTAGAGAAAGAC 1112
|||||
Db 2450 AAGACATGTGTGTAT-AGAGACTCCTCAATGATTTAGACAACCTTCAAAATACAGAGAA 2508
|||||
QY 1113 AAGACATGTGTGTATTAGAGACTCCTGAATGATTTAGACAACCTTCAAAATACAGAGAA 1172
|||||
Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATATACATTTTAAGGGGAAAAAATAA 2565
|||||
QY 1173 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATATACATTTTAAGGGGAAAAAATAA 1232
|||||
Db 2566 CCCACCATTCCTCTCCCTTATTAATTTGCAACAATAAAGGGTGGAGGGTAACTTC 2625
|||||
QY 1233 CCCACCATTCCTCTCCCTTATTAATTTGCAACAATAAAGGGTGGAGGGTAACTTC 1292
|||||
Db 2626 TACTTTCCTATACCTGCCAAAGAAATGAGGAAGAAATGGACTCTTTGCTTATTTATGA 2685
|||||
QY 1293 TACTTTCCTATACCTGCCAAAGAAATGAGGAAGAAATGGACTCTTTGCTTATTTATGA 1352
|||||
Db 2686 TGGCACTGTAATTTGTACAGTATTTCTGGAGGGCAATTTGCTAAATGCATCAAAAGAC 2745
|||||
QY 1353 TGGCACTGTAATTTGTACAGTATTTCTGGAGGGCAATTTGCTAAATGCATCAAAAGAC 1412
|||||
Db 2746 TTAATAATACGACGT 2761
|||||
QY 1413 TTAATAATACGACGT 1428
RESULT 3
ID PCT-US94-14920-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application PC/TUS9414920.
CC Sequence 1, Application PC/TUS9414920
CC GENERAL INFORMATION:
CC APPLICANT: THE REGENTS OF THE UNIVERSITY

```

CC      APPLICANT:  OF CALIFORNIA
CC      TITLE OF INVENTION:  METHOD FOR DETECTION OF
CC      TITLE OF INVENTION:  METHYLRHIOADENOSINE PHOSPHATASE DEFICIENCY IN MAMM
CC      ALIAN
CC      TITLE OF INVENTION:  CELLS
CC      NUMBER OF SEQUENCES:  5
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE:  Robbins, Berliner & Carson
CC      STREET:  201 N. Figueroa Street, 5th Floor
CC      CITY:  Los Angeles
CC      STATE:  California
CC      COUNTRY:  USA
CC      ZIP:  90012
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE:  Floppy disk
CC      COMPUTER:  IBM PC compatible
CC      OPERATING SYSTEM:  PC-DOS/MS-DOS
CC      SOFTWARE:  PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER:  PCT/US94/14920
CC      FILING DATE:
CC      CLASSIFICATION:
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME:  Berliner, Robert
CC      REGISTRATION NUMBER:  20,121
CC      REFERENCE/DOCKET NUMBER:  5555-287
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE:  213-977-1001
CC      TELEFAX:  213-977-1003
CC      INFORMATION FOR SEQ ID NO:  1:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH:  2763 base pairs
CC      TYPE:  nucleic acid
CC      STRANDEDNESS:  single
CC      TOPOLOGY:  linear
CC      MOLECULE TYPE:  DNA (genomic)
CC      IMMEDIATE SOURCE:
CC      CLONE:  methyladenosine phosphatase
CC      FEATURE:
CC      NAME/KEY:  CDS
CC      LOCATION:  1..2763
CC      Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match          20.5%; Score 466; DB 10; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db      2270 AGAATATGCCCCAGTTTCTGTTTATTATACCAAGACATTAAAGTAGCATGGCTGCCAGG 2329
        |||||||
        933 AGATATAGCCCCAGTTTCTGTTTATTATACCAAGACATTAAAGTAGCATGGCTGCCAGG 992
Db      2330 AGAAAGAAGACATTCTAATTCACGTCATTTTGGGAATTCCTGCTTAACTGAAAAAAT 2389
        |||||||
        993 AGAAAGAAGACATTCTAATTCACGTCATTTTGGGAATTCCTGCTTAACTGAAAAAAT 1052
Db      2390 ATGGGAAAGACATGACGCTTTCATGCCCCCTGCTATCAAAAGAGTATGTTAAGAAAGAC 2449
        |||||||
        1053 ATGGGAAAGACATGACGCTTTCATGCCCCCTGCTATCAAAAGAGTATGTTAAGAAAGAC 1112
Db      2450 AAGACATTGTGTAT -AGAGACTCCTCAATGATTTAGACAACCTCAAAATACAGAAGAA 2508
        |||||||
        1113 AAGACATTGTGTATAGAGACTCCTGAATGATTTAGACAACCTCAAAATACAGAAGAA 1172
Db      2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATTATCATTTTAAAGGGGGAAAAA-- 2565
        |||||||
        1173 AAGCAATGACTAGTAAACATGTGGGAAAAAATATTATCATTTTAAAGGGGGAAAAA 1232
Db      2566 CCCACCATCTCTCTCTCCCCCTATTAATTGCAACAATAAAGGTGAGGGTAAATCTC 2625
        |||||||
        1233 CCCACCATCTCTCTCTCCCCCTATTAATTGCAACAATAAAGGTGAGGGTAAATCTC 1292
Db      2626 TACTTCTTACTACTGCCAAAGAATGAGGAAGAAATGGAGCTCTTGGTTATTATTGA 2685
        |||||||

```

[illegible]

CP 1022 AATGACTGGAATTAAGATGCTTCTTTCTCCCTGGGACGCCATGCTAATTAAATGCTGTG 963
Db 125 TAMTSRNRTGKTANNVDSRNMGDASVSDKNTKKHAKNSADGKYGSKNNGDRNRNYGTG 184
CP 962 GTATATAACAGAAAGCTGGCCATATCTTCAGAGTTATGAGGGTTTCTGACCATTTCTG 903
Db 185 TKSNS 190
CP 902 TGGACC 897

RESULT 5
ID US-08-238-163-5 STANDARD; DNA; UNC; 215 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5, Application US/08238163.
CC Sequence 5, Application US/08238163.
CC Patent No. 5569830
CC GENERAL INFORMATION:
CC APPLICANT: BENNETT, Alan
CC APPLICANT: LABAVITCH, John M.
CC APPLICANT: POWELL, Ann
CC APPLICANT: STOTZ, Henrik
CC TITLE OF INVENTION: PLANT INHIBITORS OF FUNGAL
CC TITLE OF INVENTION: POLYGALACTURONASES AND THEIR USE TO CONTROL FUNGAL
CC DISEASE
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Townsend and Townsend Knourle and Crew
CC STREET: Stuart Street Tower, One Market Plaza
CC CITY: San Francisco
CC STATE: California
CC COUNTRY: US
CC ZIP: 94105-1493
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/238,163
CC FILING DATE: 03-MAY-1994
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Bastian, Kevin L.
CC REGISTRATION NUMBER: 34,774
CC REFERENCE/DOCKET NUMBER: 2307E-540
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 543-9600
CC TELEFAX: (415) 543-5043
CC INFORMATION FOR SEQ ID NO: 5:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 215 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC FEATURE:
CC NAME/KEY: misc-feature
CC LOCATION: 1..215
CC OTHER INFORMATION: /standard_name="Deduced amino acid
CC OTHER INFORMATION: sequence of PGIP from bean."
CC Sequence 215 BP; 15 A; 8 C; 25 G; 26 T; 141 other;

Query Match 1.4%; Score 31; DB 5; Length 215;
Best Local Similarity 11.98; Pred. No. 7.16e-05;
Matches 20; Conservative 72; Mismatches 75; Indels 1; Gaps 1;

Db 6 MSSSVSRTASCNDKARKDNTTSSWTDCCNRTWGVCDTDTTYRVNNDGHNKYSAN 65
OY 658 CCGGCGAAGAAAGCTTCATGTTCGACACCTGGGGGGCGGATGTATCAACATGACACACACT 717

Db 66 YNYGNVGAAKTHYHTNVSGADSKVTDSYNASGT-SSNGTDCNRSAGDSYSSK 124
OY 718 TCAGAGGTGTTCTTCTTAAGAGAGCTGAATTTGTACGCAAGTATGCGCATGCGGAC 777
Db 125 TAMTSRNRTGKTANNVDSRNMGDASVSDKNTKKHAKNSADGKYGSK 172
OY 778 AGATTATGACTGCTGGAAGAGACACAGAGAACAGATTTCGCTGACCG 825

RESULT 6
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 100, Application PC/TUS9511934.
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Ablides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC Sequence 74 BP; 6 A; 6 C; 1 G; 1 T; 60 other;

Query Match 1.1%; Score 24; DB 11; Length 74;
Best Local Similarity 6.5%; Pred. No. 2.80e-01;
Matches 4; Conservative 20; Mismatches 38; Indels 0; Gaps 0;

Db 2 TAGAVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 61
OY 1148 TAGACAACTTCAAAATACAGAGAAAGAAAGCAATGACTAGTAACATGTGGGAAAAATAT 1207

Db 62 NV 63
OY 1208 TA 1209

RESULT 7
ID PCT-US95-11934-100 STANDARD; DNA; UNC; 74 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 100, Application PC/TUS9511934.
CC Sequence 100, Application PC/TUS9511934
CC GENERAL INFORMATION:

CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 100:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 74 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 74 BP; 6 A; 6 C; 1 G; 1 T; 60 other;
SQ

Query Match 1.1%; Score 25; DB 11; Length 74;
Best Local Similarity 9.0%; Pred. No. 9.13e-02;
Matches 6; Conservative 20; Mismatches 41; Indels 0; Gaps 0;

Db 6 VNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNNVNN 65
Cp 469 CCGTCAATGACTGATCAATATGACATATCGCGCGCTGATCTCTCCCTCAAGA 410
Db 66 ACCACAA 72
Cp 409 GCCACAA 403

RESULT 8
ID PCT-US95-11934-99 STANDARD; DNA; UNC; 75 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 99, Application PC/TUS9511934.
CC Sequence 99, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 99:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 75 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC SEQUENCE 75 BP; 1 A; 1 C; 7 G; 5 T; 61 other;
SQ

Query Match 1.1%; Score 25; DB 11; Length 75;
Best Local Similarity 12.5%; Pred. No. 9.13e-02;
Matches 9; Conservative 19; Mismatches 44; Indels 0; Gaps 0;

Db 4 AGNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNBNNB 63
Cp 211 AGTCTTCCTCTTAATTTCTGATCATCCAGCGCTGTTCCACCAATATTCATCTT 152
Db 64 NNBGTTGTGT 75
Cp 151 CACGGCGGTGT 140

RESULT 9
ID PCT-US95-11934-98 STANDARD; DNA; UNC; 81 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 98, Application PC/TUS9511934.
CC Sequence 98, Application PC/TUS9511934
CC GENERAL INFORMATION:
CC APPLICANT: Cytogen Corporation
CC TITLE OF INVENTION: Antigen Binding Peptides (Abtides) From
CC TITLE OF INVENTION: Peptide Libraries
CC NUMBER OF SEQUENCES: 103
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Pennie & Edmonds
CC STREET: 1155 Avenue of the Americas
CC CITY: New York
CC STATE: New York
CC COUNTRY: USA
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US95/11934
CC FILING DATE: 20-SEP-1995
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Mistrock, S. Leslie
CC REGISTRATION NUMBER: 18,872
CC REFERENCE/DOCKET NUMBER: 1101-196-228
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 790-9090
CC TELEFAX: (212) 869-9741/8864
CC TELEX: 66141 PENNIE
CC INFORMATION FOR SEQ ID NO: 98:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 81 base pairs
CC TYPE: nucleic acid

CC LENGTH: 3214 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: CDNA
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 220..2802
SQ Sequence 3214 BP; 897 A; 808 C; 792 G; 717 T; 0 other;

Query Match 1.1%; Score 24; DB 5; Length 3214;
Best Local Similarity 100.0%; Pred. No. 2.80e-01;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3144 TTGCTTTTAAAAA 3167
QY 2237 TTGCTTTTAAAAA 2260

RESULT 15
ID PCT-US96-08014-269 STANDARD; DNA; UNC; 26 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 269, Application PC/TUS9608014.
CC Sequence 269, Application PC/TUS9608014
CC GENERAL INFORMATION:
CC APPLICANT: LARRY GOLD; NEBOUSA JANJIC; STEVEN RINGQUIST; NIKOS
CC APPLICANT: PAGRATIS; PENELOPE J. TOOTHMAN
CC TITLE OF INVENTION: HIGH AFFINITY OLIGONUCLEOTIDE
CC TITLE OF INVENTION: LIGANDS TO TRANSFORMING GROWTH
CC TITLE OF INVENTION: FACTOR (TGF), PLATELET-DERIVED
CC TITLE OF INVENTION: GROWTH FACTOR (PDGF) AND HUMAN
CC TITLE OF INVENTION: KEROTINOCYTE GROWTH FACTOR (hKGF)
CC NUMBER OF SEQUENCES: 304
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Swanson & Bratschun, L.L.C.
CC STREET: 8400 E. Prentice Avenue, Suite 200
CC CITY: Englewood
CC STATE: Colorado
CC COUNTRY: USA
CC ZIP: 80111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.5 inch, 1.44 MB storage
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: MS-DOS
CC SOFTWARE: WordPerfect 6.1
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US96/08014
CC FILING DATE:
CC CLASSIFICATION:
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/458,423
CC FILING DATE: 02-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/458,424
CC FILING DATE: 02-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/465,594
CC FILING DATE: 05-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/465,591
CC FILING DATE: 05-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/479,725
CC FILING DATE: 07-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/479,783
CC FILING DATE: 07-JUNE-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/618,693
CC FILING DATE: 20-MARCH-1996
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Barry J. Swanson

CC REGISTRATION NUMBER: 33,215
CC REFERENCE/DOCKET NUMBER:
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (303) 793-3333
CC TELEFAX: (303) 793-3433
CC INFORMATION FOR SEQ ID NO: 269:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: RNA
CC FEATURE:
CC OTHER INFORMATION: All pyrimidines are 2'-fluoro (2'-F) modified
SQ Sequence 26 BP; 1 A; 0 C; 3 G; 0 T; 22 other;

Query Match 1.0%; Score 22; DB 12; Length 26;
Best Local Similarity 16.7%; Pred. No. 2.45e+00;
Matches 4; Conservative 19; Mismatches 1; Indels 0; Gaps 0;

Db 1 RGRGRGRAYRSBSYRYBSYB 24
QY 1320 AGGAAGAAATGGACTTTGGTT 1343

Search completed: Mon Feb 10 23:15:58 1997
Job time : 91 secs.

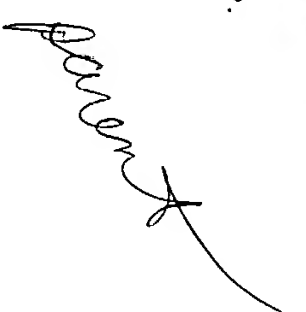
CC NAME: Kitchell, Barbara S.
CC REGISTRATION NUMBER: 33,928
CC REFERENCE/DOCKET NUMBER: ARSB:509
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (512) 474-7577
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2269 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 122..970
SQ Sequence 2269 BP; 725 A; 407 C; 490 G; 647 T; 0 other;

Query Match 100.0%; Score 2269; DB 66; Length 2269;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 GAATTCGGCTCCGCACTGCTCATCTCCCGCGCAGTGAAGTTGGCACAGCCACCGCTGTGTG 60
1 GAATTCGGCTCCGCACTGCTCATCTCCCGCGCAGTGAAGTTGGCACAGCCACCGCTGTGTG 60
QY 1 GAATTCGGCTCCGCACTGCTCATCTCCCGCGCAGTGAAGTTGGCACAGCCACCGCTGTGTG 60
Db 61 GCTCGCTTGGTCCCTTAGTCCCGAGCGCTGCCCACTGCAGATTCCTTCCCGTGCAGA 120
61 GCTCGCTTGGTCCCTTAGTCCCGAGCGCTGCCCACTGCAGATTCCTTCCCGTGCAGA 120
QY 61 GCTCGCTTGGTCCCTTAGTCCCGAGCGCTGCCCACTGCAGATTCCTTCCCGTGCAGA 120
61 GCTCGCTTGGTCCCTTAGTCCCGAGCGCTGCCCACTGCAGATTCCTTCCCGTGCAGA 120
Db 121 CATGGCCTGTGGCACCACCACCGCGGTGAAGATTGGAATTAATGGTGGACAGGCGCT 180
121 CATGGCCTGTGGCACCACCACCGCGGTGAAGATTGGAATTAATGGTGGACAGGCGCT 180
QY 121 CATGGCCTGTGGCACCACCACCGCGGTGAAGATTGGAATTAATGGTGGACAGGCGCT 180
121 CATGGCCTGTGGCACCACCACCGCGGTGAAGATTGGAATTAATGGTGGACAGGCGCT 180
Db 181 GGATGATCCAGAAATTTTAGAAGAGAAGAACTGAATAATATGTGATACCTCCATTGGCAA 240
181 GGATGATCCAGAAATTTTAGAAGAGAAGAACTGAATAATATGTGATACCTCCATTGGCAA 240
QY 181 GGATGATCCAGAAATTTTAGAAGAGAAGAACTGAATAATATGTGATACCTCCATTGGCAA 240
181 GGATGATCCAGAAATTTTAGAAGAGAAGAACTGAATAATATGTGATACCTCCATTGGCAA 240
Db 241 GCCATCTGATGCCCTTAATTTGGGGAAGATAAAAAATGTTGATTCATCTCCTTGCAAG 300
241 GCCATCTGATGCCCTTAATTTGGGGAAGATAAAAAATGTTGATTCATCTCCTTGCAAG 300
QY 241 GCCATCTGATGCCCTTAATTTGGGGAAGATAAAAAATGTTGATTCATCTCCTTGCAAG 300
241 GCCATCTGATGCCCTTAATTTGGGGAAGATAAAAAATGTTGATTCATCTCCTTGCAAG 300
Db 301 GCATGGAAGGCAGACACCATCATCTCTTCAAAAGGTCAACTACGAGCGCAACATCTGGGC 360
301 GCATGGAAGGCAGACACCATCATCTCTTCAAAAGGTCAACTACGAGCGCAACATCTGGGC 360
QY 301 GCATGGAAGGCAGACACCATCATCTCTTCAAAAGGTCAACTACGAGCGCAACATCTGGGC 360
301 GCATGGAAGGCAGACACCATCATCTCTTCAAAAGGTCAACTACGAGCGCAACATCTGGGC 360
Db 361 TTTGAAGGAAGAGGCTGTACACATGTCTATAGTACACACAGCTTGTGGCTCCTTGAGGGA 420
361 TTTGAAGGAAGAGGCTGTACACATGTCTATAGTACACACAGCTTGTGGCTCCTTGAGGGA 420
QY 361 TTTGAAGGAAGAGGCTGTACACATGTCTATAGTACACACAGCTTGTGGCTCCTTGAGGGA 420
361 TTTGAAGGAAGAGGCTGTACACATGTCTATAGTACACACAGCTTGTGGCTCCTTGAGGGA 420
Db 421 GGAGATTACAGCCCGGAGATATTGTCAATTATGATCAGTTTACATGACAGACCACTATGAG 480
421 GGAGATTACAGCCCGGAGATATTGTCAATTATGATCAGTTTACATGACAGACCACTATGAG 480
QY 421 GGAGATTACAGCCCGGAGATATTGTCAATTATGATCAGTTTACATGACAGACCACTATGAG 480
421 GGAGATTACAGCCCGGAGATATTGTCAATTATGATCAGTTTACATGACAGACCACTATGAG 480
Db 481 ACCTCAGTCTTCTATGATGGAAGTCAATCTTGTGCCAGAGAGAGTGTGCATATTCCAAT 540
481 ACCTCAGTCTTCTATGATGGAAGTCAATCTTGTGCCAGAGAGAGTGTGCATATTCCAAT 540
QY 481 ACCTCAGTCTTCTATGATGGAAGTCAATCTTGTGCCAGAGAGAGTGTGCATATTCCAAT 540
481 ACCTCAGTCTTCTATGATGGAAGTCAATCTTGTGCCAGAGAGAGTGTGCATATTCCAAT 540
Db 541 GCGTGAGCCGTTTGGCCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAGCTAGG 600
541 GCGTGAGCCGTTTGGCCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAGCTAGG 600
QY 541 GCGTGAGCCGTTTGGCCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAGCTAGG 600
541 GCGTGAGCCGTTTGGCCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAGCTAGG 600
Db 601 ACTCCGCTGCCACTCAAGAGGGGACATGTCACATCGAAGGACCTCGTTTTAGCTCCCG 660
601 ACTCCGCTGCCACTCAAGAGGGGACATGTCACATCGAAGGACCTCGTTTTAGCTCCCG 660
QY 601 ACTCCGCTGCCACTCAAGAGGGGACATGTCACATCGAAGGACCTCGTTTTAGCTCCCG 660
601 ACTCCGCTGCCACTCAAGAGGGGACATGTCACATCGAAGGACCTCGTTTTAGCTCCCG 660
Db 661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGCGGATGTTATCAACATGACCAAGTTCC 720
661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGCGGATGTTATCAACATGACCAAGTTCC 720
QY 661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGCGGATGTTATCAACATGACCAAGTTCC 720
661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGCGGATGTTATCAACATGACCAAGTTCC 720
Db 721 AGAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA 780
721 AGAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA 780
QY 721 AGAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA 780
721 AGAGTGTCTTCTGCTAAGAGGCTGGAATTTGTTACGCAAGTATCGCCATGGCGACAGA 780

Db 781 TTATGACTGCTGGAAGAGCAGCAGAGAACAGTTCGCGTGACCGCGGCTTTAAAGACCT 840
781 TTATGACTGCTGGAAGAGCAGCAGAGAACAGTTCGCGTGACCGCGGCTTTAAAGACCT 840
QY 781 TTATGACTGCTGGAAGAGCAGCAGAGAACAGTTCGCGTGACCGCGGCTTTAAAGACCT 840
781 TTATGACTGCTGGAAGAGCAGCAGAGAACAGTTCGCGTGACCGCGGCTTTAAAGACCT 840
Db 841 GAAAGAAAACGCTTAATAAGCCAAAAGCTTACTGCTCACTACCATACCTCAGATAGGCTC 900
841 GAAAGAAAACGCTTAATAAGCCAAAAGCTTACTGCTCACTACCATACCTCAGATAGGCTC 900
QY 841 GAAAGAAAACGCTTAATAAGCCAAAAGCTTACTGCTCACTACCATACCTCAGATAGGCTC 900
841 GAAAGAAAACGCTTAATAAGCCAAAAGCTTACTGCTCACTACCATACCTCAGATAGGCTC 900
Db 901 CACAGATGCTCAGAAACCCCTCCATACCTGAAGATATGGCCAGTTTCTGTTTATT 960
901 CACAGATGCTCAGAAACCCCTCCATACCTGAAGATATGGCCAGTTTCTGTTTATT 960
QY 901 CACAGATGCTCAGAAACCCCTCCATACCTGAAGATATGGCCAGTTTCTGTTTATT 960
901 CACAGATGCTCAGAAACCCCTCCATACCTGAAGATATGGCCAGTTTCTGTTTATT 960
Db 961 ACCAAGACATTTAAGTAGCATGGCTGCCAGAGAAAAGAAAGACATTTCTAATCCAGTCA 1020
961 ACCAAGACATTTAAGTAGCATGGCTGCCAGAGAAAAGAAAGACATTTCTAATCCAGTCA 1020
QY 961 ACCAAGACATTTAAGTAGCATGGCTGCCAGAGAAAAGAAAGACATTTCTAATCCAGTCA 1020
961 ACCAAGACATTTAAGTAGCATGGCTGCCAGAGAAAAGAAAGACATTTCTAATCCAGTCA 1020
Db 1021 TTTGGGAATTCCTGCTTAACCTTGAAAAAATATGGGAAAGACATGCACTTTCATGCC 1080
1021 TTTGGGAATTCCTGCTTAACCTTGAAAAAATATGGGAAAGACATGCACTTTCATGCC 1080
QY 1021 TTTGGGAATTCCTGCTTAACCTTGAAAAAATATGGGAAAGACATGCACTTTCATGCC 1080
1021 TTTGGGAATTCCTGCTTAACCTTGAAAAAATATGGGAAAGACATGCACTTTCATGCC 1080
Db 1081 TTGCTATCAAGAGATGTGTGAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG 1140
1081 TTGCTATCAAGAGATGTGTGAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG 1140
QY 1081 TTGCTATCAAGAGATGTGTGAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG 1140
1081 TTGCTATCAAGAGATGTGTGAAGAAAGACAGACATTTGTGTATTAGAGACTCCTG 1140
Db 1141 AATGATTTAGACAACCTTCAAAATACAGAAAGAAAGCAATGACTAGTAACATGTGGGA 1200
1141 AATGATTTAGACAACCTTCAAAATACAGAAAGAAAGCAATGACTAGTAACATGTGGGA 1200
QY 1141 AATGATTTAGACAACCTTCAAAATACAGAAAGAAAGCAATGACTAGTAACATGTGGGA 1200
1141 AATGATTTAGACAACCTTCAAAATACAGAAAGAAAGCAATGACTAGTAACATGTGGGA 1200
Db 1201 AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATCTCTTCCCTATTAA 1260
1201 AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATCTCTTCCCTATTAA 1260
QY 1201 AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATCTCTTCCCTATTAA 1260
1201 AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATCTCTTCCCTATTAA 1260
Db 1261 ATTGCAACAATAAAGGGTGAAGGTAATCTCTACTTCTATACCTGCCAAAGAATGTGA 1320
1261 ATTGCAACAATAAAGGGTGAAGGTAATCTCTACTTCTATACCTGCCAAAGAATGTGA 1320
QY 1261 ATTGCAACAATAAAGGGTGAAGGTAATCTCTACTTCTATACCTGCCAAAGAATGTGA 1320
1261 ATTGCAACAATAAAGGGTGAAGGTAATCTCTACTTCTATACCTGCCAAAGAATGTGA 1320
Db 1321 GGAAGAAATGGGACTCTTGGTTATTATGATGCGACTGTAAATTGGTACAGTATTCT 1380
1321 GGAAGAAATGGGACTCTTGGTTATTATGATGCGACTGTAAATTGGTACAGTATTCT 1380
QY 1321 GGAAGAAATGGGACTCTTGGTTATTATGATGCGACTGTAAATTGGTACAGTATTCT 1380
1321 GGAAGAAATGGGACTCTTGGTTATTATGATGCGACTGTAAATTGGTACAGTATTCT 1380
Db 1381 GGAGGCAATTTGTAATAATGCATCAAAAGACTTAAAAATACGGAGCTGCTTGTGCTG 1440
1381 GGAGGCAATTTGTAATAATGCATCAAAAGACTTAAAAATACGGAGCTGCTTGTGCTG 1440
QY 1381 GGAGGCAATTTGTAATAATGCATCAAAAGACTTAAAAATACGGAGCTGCTTGTGCTG 1440
1381 GGAGGCAATTTGTAATAATGCATCAAAAGACTTAAAAATACGGAGCTGCTTGTGCTG 1440
Db 1441 GGAAGCTTACATCTAGCAATTTCTCTTTAAACCATATACAGATGCATACAAAGATTA 1500
1441 GGAAGCTTACATCTAGCAATTTCTCTTTAAACCATATACAGATGCATACAAAGATTA 1500
QY 1441 GGAAGCTTACATCTAGCAATTTCTCTTTAAACCATATACAGATGCATACAAAGATTA 1500
1441 GGAAGCTTACATCTAGCAATTTCTCTTTAAACCATATACAGATGCATACAAAGATTA 1500
Db 1501 TATATAAAGAGGGTGTTAATAATGATAGTTTATAATAATAATTGAACAATCTGA 1560
1501 TATATAAAGAGGGTGTTAATAATGATAGTTTATAATAATAATAATTGAACAATCTGA 1560
QY 1501 TATATAAAGAGGGTGTTAATAATGATAGTTTATAATAATAATAATTGAACAATCTGA 1560
1501 TATATAAAGAGGGTGTTAATAATGATAGTTTATAATAATAATAATTGAACAATCTGA 1560
Db 1561 ATCCCTTGCAATTTGAGGTAATGATGCTTAAATCTAGATTGTAATCAGCCAA 1620
1561 ATCCCTTGCAATTTGAGGTAATGATGCTTAAATCTAGATTGTAATCAGCCAA 1620
QY 1561 ATCCCTTGCAATTTGAGGTAATGATGCTTAAATCTAGATTGTAATCAGCCAA 1620
1561 ATCCCTTGCAATTTGAGGTAATGATGCTTAAATCTAGATTGTAATCAGCCAA 1620
Db 1621 CTGAAATCTTTTGGCATATTTCATGTCCTTAAAGACACGGTGTGCTATATATGAA 1680
1621 CTGAAATCTTTTGGCATATTTCATGTCCTTAAAGACACGGTGTGCTATATATGAA 1680
QY 1621 CTGAAATCTTTTGGCATATTTCATGTCCTTAAAGACACGGTGTGCTATATATGAA 1680
1621 CTGAAATCTTTTGGCATATTTCATGTCCTTAAAGACACGGTGTGCTATATATGAA 1680
Db 1681 GTGAAAAAAGGATATGTTAGCATTTTATAGTACTAGTTTGTCTTTAAATGCTATGTAA 1740
1681 GTGAAAAAAGGATATGTTAGCATTTTATAGTACTAGTTTGTCTTTAAATGCTATGTAA 1740
QY 1681 GTGAAAAAAGGATATGTTAGCATTTTATAGTACTAGTTTGTCTTTAAATGCTATGTAA 1740
1681 GTGAAAAAAGGATATGTTAGCATTTTATAGTACTAGTTTGTCTTTAAATGCTATGTAA 1740
Db 1741 TATACAAAAAACTAGAAAGAAATATATATACCTTGTATATGTTGGGGAGGAGATA 1800
1741 TATACAAAAAACTAGAAAGAAATATATATACCTTGTATATGTTGGGGAGGAGATA 1800
QY 1741 TATACAAAAAACTAGAAAGAAATATATATACCTTGTATATGTTGGGGAGGAGATA 1800
1741 TATACAAAAAACTAGAAAGAAATATATATACCTTGTATATGTTGGGGAGGAGATA 1800
Db 1801 CTGGGATAAATTTTATTTCTTGAATCTTCTGCTTCAATTTTCTACAGTGAAT 1860
1801 CTGGGATAAATTTTATTTCTTGAATCTTCTGCTTCAATTTTCTACAGTGAAT 1860
QY 1801 CTGGGATAAATTTTATTTCTTGAATCTTCTGCTTCAATTTTCTACAGTGAAT 1860
1801 CTGGGATAAATTTTATTTCTTGAATCTTCTGCTTCAATTTTCTACAGTGAAT 1860
Db 1861 ATAATCAAAATAGTAAAGGGCGGTAAATAAAGTGATTTAGAAAGATCCAGTCTTGA 1920
1861 ATAATCAAAATAGTAAAGGGCGGTAAATAAAGTGATTTAGAAAGATCCAGTCTTGA 1920

|||||
QY 1861 ATATCAATAGTAAAGGCCCTAAATAAAGTGATTAGAAAGATCCAGTCTTGA 1920
Db 1921 AAACACTGTTCTGTAATGAAGCAGAAATTAAGTGTATATTAAGTGAATGTCAT 1980
QY 1921 AAACACTGTTCTGTAATGAAGCAGAAATTAAGTGTATATTAAGTGAATGTCAT 1980
Db 1981 TAAGGAGTTACATCTTATCTGCTAAAGAGAGATCATGATTTCTGTACAGTCAGA 2040
QY 1981 TAAGGAGTTACATCTTATCTGCTAAAGAGAGATCATGATTTCTGTACAGTCAGA 2040
Db 2041 ACAGTACTGGGTGCAACAGCTTCTGTAGAAAAGCTAGGTATATAAGTTACTGA 2100
QY 2041 ACAGTACTGGGTGCAACAGCTTCTGTAGAAAAGCTAGGTATATAAGTTACTGA 2100
Db 2101 AAGTTACTATTTAAAGACTAAATGCACATTTATGCTATCTGATATTTTAAAGTA 2160
QY 2101 AAGTTACTATTTAAAGACTAAATGCACATTTATGCTATCTGATATTTTAAAGTA 2160
Db 2161 ATGTGAGCTTCTCTTTTATGAGTTAAATATTTATACGAGTTGGTAATTTGTGCTT 2220
QY 2161 ATGTGAGCTTCTCTTTTATGAGTTAAATATTTATACGAGTTGGTAATTTGTGCTT 2220
Db 2221 TTAATAAGTGAAGCTTCTTTTAAAAAAGCGGAATTC 2269
QY 2221 TTAATAAGTGAAGCTTCTTTTAAAAAAGCGGAATTC 2269



RESULT 2
ID US-60-000-831-1 STANDARD; DNA; UNC; 2269 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/60000831.
CC Sequence 1, Application US/60000831
CC GENERAL INFORMATION:
CC APPLICANT: Olopade, Olufunmilayo I.
CC APPLICANT: Bohlander, Stefan K.
CC APPLICANT: Dreyling, Martin H.
CC APPLICANT: Le Beau, Michelle M.
CC APPLICANT: Stadler, Walter M.
CC APPLICANT: Gursky, Susan
CC APPLICANT: Espinosa, III, Rafael
CC APPLICANT: Sveen, Lise W.
CC APPLICANT: Hagos, Filsun
CC APPLICANT: Pomykala, Helen M.
CC APPLICANT: Adeyanju, M.O.
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHORYLASE
CC TITLE OF INVENTION: METHODS AND COMPOSITIONS
CC NUMBER OF SEQUENCES: 2
CC CORRESPONDENCE ADDRESSES:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: United States of America
CC ZIP: 72210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/60/000,831
CC FILING DATE: Concurrently Herewith
CC CLASSIFICATION:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Kitchell, Barbara S.
CC REGISTRATION NUMBER: 33,928
CC REFERENCE/DOCKET NUMBER: ARCD:156/KIT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (512) 418-3000
CC TELEFAX: (512) 474-7577
CC INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
CC LENGTH: 2269 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: other nucleic acid
CC DESCRIPTION: /desc = "DNA"
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 122..970
SQ Sequence 2269 BP; 725 A; 407 C; 490 G; 647 T; 0 other;
Query Match 100.0%; Score 2269; DB 68; length 2269;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 2269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 GAATCCGCTCCGCACTGCTCACTCCCGCGAGTGAGTTGGCACAGCCACCGCTGTG 60
QY 1 GAATCCGCTCCGCACTGCTCACTCCCGCGAGTGAGTTGGCACAGCCACCGCTGTG 60
Db 61 GCTCGCTGTTCCCTTAGTCCCGAGCGCTCGCCCACTGCAGATTCCTTCCCGTCAGA 120
QY 61 GCTCGCTGTTCCCTTAGTCCCGAGCGCTCGCCCACTGCAGATTCCTTCCCGTCAGA 120
Db 121 CATGCCCTCTGGCACACACACCGCCGCTGAAGATTGGAATTAATTGGTGAACAGGCT 180
QY 121 CATGCCCTCTGGCACACACACCGCCGCTGAAGATTGGAATTAATTGGTGAACAGGCT 180
Db 181 GGATGATCCAGAAATTTTGAAGAGAACTGAATAATATGTGATGCTATCTCCATTGGCAA 240
QY 181 GGATGATCCAGAAATTTTGAAGAGAACTGAATAATATGTGATGCTATCTCCATTGGCAA 240
Db 241 GCCATCTGATGCCCTTAATTTTGGGAAGATAAATAATGTGATGCTATCTCCATTGGCAA 300
QY 241 GCCATCTGATGCCCTTAATTTTGGGAAGATAAATAATGTGATGCTATCTCCATTGGCAA 300
Db 301 GCATGGAAGGACACACCATCATGCTCTTCAAGGTCACACTACAGCGGAACATCTGGGC 360
QY 301 GCATGGAAGGACACACCATCATGCTCTTCAAGGTCACACTACAGCGGAACATCTGGGC 360
Db 361 TTTGAAGGAAGAGGGCTGTACACATGTCAATAGTGACACACAGCTTGCGCTTGAGGGA 420
QY 361 TTTGAAGGAAGAGGGCTGTACACATGTCAATAGTGACACACAGCTTGCGCTTGAGGGA 420
Db 421 GGAGATTGAGCCCGGCGATATTGTCAATTTGATGATCAAGTTCAATGACAGCACACTATGAG 480
QY 421 GGAGATTGAGCCCGGCGATATTGTCAATTTGATGATCAAGTTCAATGACAGCACACTATGAG 480
Db 481 ACCTCAGTCTCTATGATGAGAGTCAATCTTGTGCCAGAGAGAGTGTCCATATTTCCAAT 540
QY 481 ACCTCAGTCTCTATGATGAGAGTCAATCTTGTGCCAGAGAGAGTGTCCATATTTCCAAT 540
Db 541 GGCTGAGCCGTTTGGCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAAGCTAGG 600
QY 541 GGCTGAGCCGTTTGGCCCAAAACGAGAGAGGTTCTTATAGAGACTGCTAAGAAAGCTAGG 600
Db 601 ACTCCGGTGCACACTCAAGGGGCAATGTCACATCGAGGGAGCTGTTAGCTCCCG 660
QY 601 ACTCCGGTGCACACTCAAGGGGCAATGTCACATCGAGGGAGCTGTTAGCTCCCG 660
Db 661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGGATGTTATCAACATGACACAGTTCC 720
QY 661 GGCAGAAAGCTTCATGTTCCGACCTGGGGGGGATGTTATCAACATGACACAGTTCC 720
Db 721 AGAGGTGTTCTTGTGAAGGAGGCTGAATTTGTACGCAAGTATGCCATGGCGACAGA 780
QY 721 AGAGGTGTTCTTGTGAAGGAGGCTGAATTTGTACGCAAGTATGCCATGGCGACAGA 780
Db 781 TTATGACTGCTGAAGAGACAGAGGAAGCAATTTCGGTGGACCGGCTTTAAAGACCT 840
QY 781 TTATGACTGCTGAAGAGACAGAGGAAGCAATTTCGGTGGACCGGCTTTAAAGACCT 840
Db 841 GAAAGAAAACGCTAATAAAGCCAAAAGCTTACTGCTCACTACATACCTCAGATAGGGTC 900

Qy	841	GAAGAAAAACGCTAATAAAGCCAAAAGCTTACTGCTCCTACCATACCTACGATAGAGTGC	900
Db	901	CACAGAATGGTCAGAAACCCTCCATACTGAAGAATAAGGCCAGTTTCTGTTTAT	960
Qy	901	CACAGATGGTCAGAAACCCTCCATACTGAAGAATAAGGCCAGTTTCTGTTTAT	960
Db	961	ACCAAGACATTAAAGTAGCATGGCTGCCAGAGAGAAAAAGACATTCTAATTCAGTCA	1020
Qy	961	ACCAAGACATTAAAGTAGCATGGCTGCCAGAGAGAAAAAGACATTCTAATTCAGTCA	1020
Db	1021	TTTTGGGAATTCCTGCTTAACCTTGAAAAAAATATGGGAAGAAGACATGCAGCTTTCATGCC	1080
Qy	1021	TTTTGGGAATTCCTGCTTAACCTTGAAAAAAATATGGGAAGAAGACATGCAGCTTTCATGCC	1080
Db	1081	TTCCTATCAAGAGTATGTTGTAAAGAACACAGACATTGTTGTATTAGAGACTCCTG	1140
Qy	1081	TTCCTATCAAGAGTATGTTGTAAAGAACACAGACATTGTTGTATTAGAGACTCCTG	1140
Db	1141	AATGATTAGACAACTTCAAAATACAGAAAGAAAGCAAAATGACTAGTAACATGTGGAA	1200
Qy	1141	AATGATTAGACAACTTCAAAATACAGAAAGAAAGCAAAATGACTAGTAACATGTGGAA	1200
Db	1201	AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATTCTCTCTCCCTATTAA	1260
Qy	1201	AAAATATTACATTTTAAAGGGGAAAAAAACCCACCATTCTCTCTCCCTATTAA	1260
Db	1261	ATTTGCACATAAAGGCTGGAGGGTAATCTACTTTCCTATCTGCCCCAAGAGTGA	1320
Qy	1261	ATTTGCACATAAAGGCTGGAGGGTAATCTACTTTCCTATCTGCCCCAAGAGTGA	1320
Db	1321	GGAAGAAATGGGACTCTTGGTTATTATTGATCGCACTGTAATTTGGTACAGTATTCT	1380
Qy	1321	GGAAGAAATGGGACTCTTGGTTATTATTGATCGCACTGTAATTTGGTACAGTATTCT	1380
Db	1381	GGAGGGCAATTTGGTAAATGCAATCAAAAGACTTAAAAATACGAGCTCTTGGTCTG	1440
Qy	1381	GGAGGGCAATTTGGTAAATGCAATCAAAAGACTTAAAAATACGAGCTCTTGGTCTG	1440
Db	1441	GGAACCTACATCTAGCAATTTCTTTAAACCATATCAGAGATGATACAAAGAAATTA	1500
Qy	1441	GGAACCTACATCTAGCAATTTCTTTAAACCATATCAGAGATGATACAAAGAAATTA	1500
Db	1501	TATATAAAGAGGGTGTATTAAATGATAGTTATATTAATAATTAATGAACAATCTGA	1560
Qy	1501	TATATAAAGAGGGTGTATTAAATGATAGTTATATTAATAATTAATGAACAATCTGA	1560
Db	1561	ATCCCTTGCAATTTGGAGGTAATATATGCTTAGTTATATCTAGATTGGAATCAGCCAA	1620
Qy	1561	ATCCCTTGCAATTTGGAGGTAATATATGCTTAGTTATATCTAGATTGGAATCAGCCAA	1620
Db	1621	CTGAATAATCCTTTTGCATATTTCAATGTCCTAAAAAGACACGGTGTCTATATATGAA	1680
Qy	1621	CTGAATAATCCTTTTGCATATTTCAATGTCCTAAAAAGACACGGTGTCTATATATGAA	1680
Db	1681	GTAAGAAAAAGATATGTTAGCATTTTATAGTACTAGTTTGCCTTAAATGCTATGTAA	1740
Qy	1681	GTAAGAAAAAGATATGTTAGCATTTTATAGTACTAGTTTGCCTTAAATGCTATGTAA	1740
Db	1741	TATACAAAAAACTAGAAAGAAATATATAAACCCTGTTATTTGGGGAGGAGATA	1800
Qy	1741	TATACAAAAAACTAGAAAGAAATATATAAACCCTGTTATTTGGGGAGGAGATA	1800
Db	1801	CTGGATATTTTATTTCTTGAATCTTCTGTGCTTCAATTTTCTACAGTGAAT	1860
Qy	1801	CTGGATATTTTATTTCTTGAATCTTCTGTGCTTCAATTTTCTACAGTGAAT	1860
Db	1861	ATAATCAAAATAGTAAAGGGCCGTAAAAATAAAGTGATTTAGAAAGATCCAGTCTTGA	1920
Qy	1861	ATAATCAAAATAGTAAAGGGCCGTAAAAATAAAGTGATTTAGAAAGATCCAGTCTTGA	1920
Db	1921	AAACACTGTTTCTGTTATGAGCAGAAATTTAAGTTGTAATTAAGTGAATGTCATT	1980

OY	1921	AAACACTGTTTCTGTGAATGAGCAGAATTTAAGTTGGTAATAATTAAGGTGAATGCAT	1980
Db	1981	TAAAGGAGTTACATCTTTATTTCTGCTAAAGAGAGGATCATTTGTTCTGTACAGTCAGA	2040
OY	1981	TAAAGGAGTTACATCTTTATTTCTGCTAAAGAGAGGATCATTTGTTCTGTACAGTCAGA	2040
Db	2041	ACAGTACTGGGTGTGCAACAGCTTTCTGAGAAAAGCTAGGTGTATAAGTTAAGTTA	2100
OY	2041	ACAGTACTGGGTGTGCAACAGCTTTCTGAGAAAAGCTAGGTGTATAAGTTAAGTTA	2100
Db	2101	AAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA	2160
OY	2101	AAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTAAGTTA	2160
Db	2161	ATGTGAGCTTCTCTTTTATGAGTTAAATTATTTATGACGAGTTGGTAATTTTGCCCT	2220
OY	2161	ATGTGAGCTTCTCTTTTATGAGTTAAATTATTTATGACGAGTTGGTAATTTTGCCCT	2220
Db	2221	TTAATAAGTGAAGCTTGTCTTTTAAAAAAGCGGAATTC	2269
OY	2221	TTAATAAGTGAAGCTTGTCTTTTAAAAAAGCGGAATTC	2269
RESULT	3		
ID	US-08-445-648-14	STANDARD; DNA: UNC, 2763 BP.	
AC	xxxxxx		
DT	01-JAN-1900		
DE	Sequence 14, Application US/08445648.		
CC	Sequence 14, Application US/08445648		
CC	GENERAL INFORMATION:		
CC	APPLICANT: CARSON, DENNIS A.		
CC	APPLICANT: NOBORT, TSUTOMU		
CC	TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR		
CC	TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI		
ON	CANCER		
CC	TITLE OF INVENTION: TREATMENT		
CC	NUMBER OF SEQUENCES: 18		
CC	CORRESPONDENCE ADDRESS:		
CC	ADDRESSEE: Fish & Richardson P.C.		
CC	STREET: 4225 Executive Square, Suite 1400		
CC	CITY: La Jolla		
CC	STATE: California		
CC	COUNTRY: US		
CC	ZIP: 92037		
CC	COMPUTER READABLE FORM:		
CC	MEDIUM TYPE: Floppy disk		
CC	COMPUTER: IBM PC compatible		
CC	OPERATING SYSTEM: PC-DOS/MS-DOS		
CC	SOFTWARE: PatentIn Release #1.0, Version #1.25		
CC	CURRENT APPLICATION DATA:		
CC	APPLICATION NUMBER: US/08/445,648		
CC	FILING DATE: 22-May-1995		
CC	CLASSIFICATION: 435		
CC	PRIOR APPLICATION DATA:		
CC	APPLICATION NUMBER: 08/227,800		
CC	FILING DATE: 14-APR-1994		
CC	ATTORNEY/AGENT INFORMATION:		
CC	NAME: HOWELLS, STACY L.		
CC	REGISTRATION NUMBER: 34,842		
CC	REFERENCE/DOCKET NUMBER: 07340/023001		
CC	TELECOMMUNICATION INFORMATION:		
CC	TELEPHONE: 619/678-5070		
CC	TELEFAX: 619/678-5099		
CC	INFORMATION FOR SEQ ID NO: 14:		
CC	SEQUENCE CHARACTERISTICS:		
CC	LENGTH: 2763 base pairs		
CC	TYPE: nucleic acid		
CC	STRANDEDNESS: single		
CC	TOPOLOGY: linear		
CC	MOLECULE TYPE: DNA (genomic)		
CC	IMMEDIATE SOURCE:		
CC	CLONE: Methylothioadenosine Phosphorylase (genomic)		
CC	FEATURE:		

CC NAME/KEY: CDS
CC LOCATION: 1..2763
SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;

Query Match
Best Local Similarity 99.0%; Score 466; DB 59; Length 2763;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

Db 2270 AGAATATGCCCCAGTTTCTGTTTATTTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
|||
QY 933 AGAATATGCCCCAGTTTCTGTTTATTTACCAAGACATTAAGTAGCATGGCTGCCAGG 992

Db 2330 AGAAAAGAGACATTTCTAATTCAGTCAATTTGGGAATTCCTGCTTAAGTGAAGAAAAAT 2389
|||
QY 993 AGAAAAGAGACATTTCTAATTCAGTCAATTTGGGAATTCCTGCTTAAGTGAAGAAAAAT 1052

Db 2390 ATGGGAAAGACATGACGCTTTTCATGCCCCCTTGCCTATCAAGAGATAGTGTGAAGAAAGAC 2449
|||
QY 1053 ATGGGAAAGACATGACGCTTTTCATGCCCCCTTGCCTATCAAGAGATAGTGTGAAGAAAGAC 1112

Db 2450 AAGACATTTGTGTAT-AGAGACTCCTCAATGATTTAGACAACTCAAAATACAGAAGAA 2508
|||
QY 1113 AAGACATTTGTGTATAGAGACTCCTCAATGATTTAGACAACTCAAAATACAGAAGAA 1172

Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATTACATTTTAAGGGGAAAAAA-- 2565
|||
QY 1173 AAGCAATGACTAGTAAACATGTGGGAAAAAATATTACATTTTAAGGGGAAAAAA-- 1232

Db 2566 CCCACCATCTCTCTCTCCCTATTAAATTTGCAACAATAAGGGTGAAGGTAATCTC 2625
|||
QY 1233 CCCACCATCTCTCTCTCCCTATTAAATTTGCAACAATAAGGGTGAAGGTAATCTC 1292

Db 2626 TACTTTCCTATACGCCCAAGATGTGAGGAGAAATGGACTCTTGGTTATTATTTGA 2685
|||
QY 1293 TACTTTCCTATACGCCCAAGATGTGAGGAGAAATGGACTCTTGGTTATTATTTGA 1352

Db 2686 TGGGACTGTAATTTGGTACAGTATTTCTGAGGGCAATTTGGTAAATGCATCAAAAGAC 2745
|||
QY 1353 TGGGACTGTAATTTGGTACAGTATTTCTGAGGGCAATTTGGTAAATGCATCAAAAGAC 1412

Db 2746 TTAAAAATACGGAGCT 2761
|||
QY 1413 TTAAAAATACGGAGCT 1428

RESULT 4
ID US-08-176-855-1 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 1, Application US/08176855.
CC Sequence 1, Application US/08176855.
CC GENERAL INFORMATION:
CC APPLICANT: Nobori, Tsutomu
CC APPLICANT: Carson, Dennis A.
CC APPLICANT: Takabayashi, Kenji
CC TITLE OF INVENTION: METHOD FOR DETECTION OF
CC TITLE OF INVENTION: METHYLTHIOADENOSINE PHOSPHATASE DEFICIENCY IN MAMM
CC ALIAN
CC TITLE OF INVENTION: CELLS
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Spensley Horn Jubas & Lubitz
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: USA
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/176,855
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Howells, Stacy L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD3057
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (619) 455-5100
CC TELEFAX: (619) 455-5110
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methyladenosine phosphatase
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
SQ Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;

Query Match
Best Local Similarity 99.0%; Score 466; DB 32; Length 2763;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;

Db 2270 AGAATATGCCCCAGTTTCTGTTTATTTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
|||
QY 933 AGAATATGCCCCAGTTTCTGTTTATTTACCAAGACATTAAGTAGCATGGCTGCCAGG 992

Db 2330 AGAAAAGAGACATTTCTAATTCAGTCAATTTGGGAATTCCTGCTTAAGTGAAGAAAAAT 2389
|||
QY 993 AGAAAAGAGACATTTCTAATTCAGTCAATTTGGGAATTCCTGCTTAAGTGAAGAAAAAT 1052

Db 2390 ATGGGAAAGACATGACGCTTTTCATGCCCCCTTGCCTATCAAGAGATAGTGTGAAGAAAGAC 2449
|||
QY 1053 ATGGGAAAGACATGACGCTTTTCATGCCCCCTTGCCTATCAAGAGATAGTGTGAAGAAAGAC 1112

Db 2450 AAGACATTTGTGTAT-AGAGACTCCTCAATGATTTAGACAACTCAAAATACAGAAGAA 2508
|||
QY 1113 AAGACATTTGTGTATAGAGACTCCTCAATGATTTAGACAACTCAAAATACAGAAGAA 1172

Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATTACATTTTAAGGGGAAAAAA-- 2565
|||
QY 1173 AAGCAATGACTAGTAAACATGTGGGAAAAAATATTACATTTTAAGGGGAAAAAA-- 1232

Db 2566 CCCACCATCTCTCTCTCCCTATTAAATTTGCAACAATAAGGGTGAAGGTAATCTC 2625
|||
QY 1233 CCCACCATCTCTCTCTCCCTATTAAATTTGCAACAATAAGGGTGAAGGTAATCTC 1292

Db 2626 TACTTTCCTATACGCCCAAGATGTGAGGAGAAATGGACTCTTGGTTATTATTTGA 2685
|||
QY 1293 TACTTTCCTATACGCCCAAGATGTGAGGAGAAATGGACTCTTGGTTATTATTTGA 1352

Db 2686 TGGGACTGTAATTTGGTACAGTATTTCTGAGGGCAATTTGGTAAATGCATCAAAAGAC 2745
|||
QY 1353 TGGGACTGTAATTTGGTACAGTATTTCTGAGGGCAATTTGGTAAATGCATCAAAAGAC 1412

Db 2746 TTAAAAATACGGAGCT 2761
|||
QY 1413 TTAAAAATACGGAGCT 1428

RESULT 5
ID US-08-443-880-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08443880.
CC Sequence 14, Application US/08443880.
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.

CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
CC ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/443,880
CC FILING DATE: 22-MAY-1995
CC CLASSIFICATION: 530
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: 08/227,800
CC FILING DATE: 14-APR-1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 20.5%; Score 466; DB 59; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
| | | | |
QY 933 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 992
Db 2330 AGAAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGAAAT 2389
| | | | |
QY 993 AGAAAAGAGACATTTCTAATTCAGTCATTTTGGGAATTCCTGCTTAAGTGAAGAAAT 1052
Db 2390 ATGGGAAGACATGCAGCTTTCATGCCCCCTGCTATCAAGAGATATGTTGTAAGAAAGAC 2449
| | | | |
QY 1053 ATGGGAAGACATGCAGCTTTCATGCCCCCTGCTATCAAGAGATATGTTGTAAGAAAGAC 1112
Db 2450 AAGACATTTGTTGTTAT-AGAGACTCCCTCAATGATTTAGACAACCTCAAAATACAGAGAA 2508
| | | | |
QY 1113 AAGACATTTGTTGTTATAGAGACTCCCTCAATGATTTAGACAACCTCAAAATACAGAGAA 1172
Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAAATATTACATTTTAAGGGGGGAAAAAATAA 2565
| | | | |
QY 1173 AAGCAATGACTAGTAAACATGTGGGAAAAAATATTACATTTTAAGGGGGGAAAAAATAA 1232
| | | | |
Db 2566 CCCACCATTTCTTCTCCCTATTAAATTTGACACAATAAAGGTTGAGAGGTAATCTC 2625
| | | | |
QY 1233 CCCACCATTTCTTCTCCCTATTAAATTTGACACAATAAAGGTTGAGAGGTAATCTC 1292

Db 2626 TACTTCCCTATAGTCCCAAGAAATGTGAGGAAGAAATGGGACTCTTGGTTATTATTGA 2685
| | | | |
QY 1293 TACTTCCCTATAGTCCCAAGAAATGTGAGGAAGAAATGGGACTCTTGGTTATTATTGA 1352
Db 2686 TGGCAGCTGAATTTGTATACAGTATTCTGAGGGCAATTTGTTAAATGCATCAAAAGAC 2745
| | | | |
QY 1353 TGGCAGCTGAATTTGTATACAGTATTCTGAGGGCAATTTGTTAAATGCATCAAAAGAC 1412
Db 2746 TTAATAATACGGACGT 2761
| | | | |
QY 1413 TTAATAATACGGACGT 1428
RESULT 6
ID US-08-227-800A-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800A.
CC Sequence 14, Application US/08227800A
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC APPLICANT: NOBORI, TSUTOMU
CC TITLE OF INVENTION: TUMOR SUPPRESSOR GENE AND METHODS FOR
CC TITLE OF INVENTION: DETECTION OF CANCER, MONITORING OF TUMOR PROGRESSI
CC ON AND CANCER
CC TITLE OF INVENTION: TREATMENT
CC NUMBER OF SEQUENCES: 18
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Fish & Richardson P.C.
CC STREET: 4225 Executive Square, Suite 1400
CC CITY: La Jolla
CC STATE: California
CC COUNTRY: US
CC ZIP: 92037
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800A
CC FILING DATE: 14-APR-1994
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: 07340/023001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/678-5070
CC TELEFAX: 619/678-5099
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: Methylthioadenosine Phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 20.5%; Score 466; DB 37; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 2329
| | | | |
QY 933 AGAATATGCCCCAGTTTCTGTTTATTACCAAGACATTAAGTAGCATGGCTGCCAGG 992

Db 2330 AGAAAGAAGACATCTAATTCAGTATTTGGGAATTCCTGCTTAAGTGAAGAAAT 2389
|||||
Qy 993 AGAAAGAAGACATCTAATTCAGTATTTGGGAATTCCTGCTTAAGTGAAGAAAT 1052
Db 2390 ATGGGAAGACATGACCTTTTCATGCCCCCTTATCAAGAGATGTTGTAAGAAAGAC 2449
|||||
Qy 1053 ATGGGAAGACATGACCTTTTCATGCCCCCTTATCAAGAGATGTTGTAAGAAAGAC 1112
Db 2450 AAGACATGTTGTAT-AGAGACTCCTCAATGATTAGACAACCTCAAAATACAGAAGAA 2508
|||||
Qy 1113 AAGACATGTTGTATAGAGACTCCTGAATGATTAGACAACCTCAAAATACAGAAGAA 1172
Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAATATTACATTTTAAAGGGGAAAAAA-- 2565
|||||
Qy 1173 AAGCAATGACTAGTAAACATGTGGGAAAAATATTACATTTTAAAGGGGAAAAAA 1232
Db 2566 CCCACCATCTCTCTCCCTATTAAATTGCAACAATAAGGGTGAAGGTAATCTC 2625
|||||
Qy 1233 CCCACCATCTCTCTCCCTATTAAATTGCAACAATAAGGGTGAAGGTAATCTC 1292
Db 2626 TACTTCTATACGCAAGAATGTGAGAGAAGATGGACTCTTGTATTATTGA 2685
|||||
Qy 1293 TACTTCTATACGCAAGAATGTGAGAGAAGATGGACTCTTGTATTATTGA 1352
Db 2686 TCGGACTGTAATGTGTACAGTATTTCTGGAGGGCAATTTGGTAAATGCATCAAGAAC 2745
|||||
Qy 1353 TCGGACTGTAATGTGTACAGTATTTCTGGAGGGCAATTTGGTAAATGCATCAAGAAC 1412
Db 2746 TTAATAATACGACGT 2761
|||||
Qy 1413 TTAATAATACGACGT 1428

RESULT 7

ID US-08-227-800-14 STANDARD; DNA; UNC; 2763 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08227800.
CC Sequence 14, Application US/08227800
CC GENERAL INFORMATION:
CC APPLICANT: CARSON, DENNIS A.
CC TITLE OF INVENTION: NOBORI, TSUTOMU
CC TITLE OF INVENTION: METHODS FOR TREATMENT AND DETECTION OF
CC TITLE OF INVENTION: SUSCEPTIBILITY TO, AND PROGRESSION OF, CANCER
CC NUMBER OF SEQUENCES: 14
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
CC STREET: 1880 Century Park East, Suite 500
CC CITY: Los Angeles
CC STATE: California
CC COUNTRY: US
CC ZIP: 90067
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/227,800
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: HOWELLS, STACY L.
CC REGISTRATION NUMBER: 34,842
CC REFERENCE/DOCKET NUMBER: PD3459
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 619/455-5110
CC TELEFAX: 619/455-5110
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2763 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single

CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC IMMEDIATE SOURCE:
CC CLONE: methylthioadenosine phosphorylase (genomic)
CC FEATURE:
CC NAME/KEY: CDS
CC LOCATION: 1..2763
CC Sequence 2763 BP; 780 A; 505 C; 598 G; 803 T; 77 other;
SQ
Query Match 20.5%; Score 466; DB 37; Length 2763;
Best Local Similarity 99.0%; Pred. No. 0.00e+00;
Matches 491; Conservative 0; Mismatches 1; Indels 4; Gaps 3;
Db 2270 AGAATATGCCCCGTTTCTGTTTATTACCAAGACATTAAAGTAGCATGGCTGCCAGG 2329
|||||
Qy 933 AGAATATGCCCCGTTTCTGTTTATTACCAAGACATTAAAGTAGCATGGCTGCCAGG 992
Db 2330 AGAAAGAAGACATCTAATTCAGTATTTGGGAATTCCTGCTTAAGTGAAGAAAT 2389
|||||
Qy 993 AGAAAGAAGACATCTAATTCAGTATTTGGGAATTCCTGCTTAAGTGAAGAAAT 1052
Db 2390 ATGGGAAGACATGACCTTTTCATGCCCCCTTATCAAGAGATGTTGTAAGAAAGAC 2449
|||||
Qy 1053 ATGGGAAGACATGACCTTTTCATGCCCCCTTATCAAGAGATGTTGTAAGAAAGAC 1112
Db 2450 AAGACATGTTGTAT-AGAGACTCCTCAATGATTAGACAACCTCAAAATACAGAAGAA 2508
|||||
Qy 1113 AAGACATGTTGTATAGAGACTCCTGAATGATTAGACAACCTCAAAATACAGAAGAA 1172
Db 2509 AAGCAATGACTAGTAA-CATGTGGGAAAAATATTACATTTTAAAGGGGAAAAAA-- 2565
|||||
Qy 1173 AAGCAATGACTAGTAAACATGTGGGAAAAATATTACATTTTAAAGGGGAAAAAA 1232
Db 2566 CCCACCATCTCTCTCCCTATTAAATTGCAACAATAAGGGTGAAGGTAATCTC 2625
|||||
Qy 1233 CCCACCATCTCTCTCCCTATTAAATTGCAACAATAAGGGTGAAGGTAATCTC 1292
Db 2626 TACTTCTATACGCAAGAATGTGAGAGAAGATGGACTCTTGTATTATTGA 2685
|||||
Qy 1293 TACTTCTATACGCAAGAATGTGAGAGAAGATGGACTCTTGTATTATTGA 1352
Db 2686 TCGGACTGTAATGTGTACAGTATTTCTGGAGGGCAATTTGGTAAATGCATCAAGAAC 2745
|||||
Qy 1353 TCGGACTGTAATGTGTACAGTATTTCTGGAGGGCAATTTGGTAAATGCATCAAGAAC 1412
Db 2746 TTAATAATACGACGT 2761
|||||
Qy 1413 TTAATAATACGACGT 1428

RESULT 8

ID US-08-220-691-5097 STANDARD; DNA; UNC; 495 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5097, Application US/08220691.
CC Sequence 5097, Application US/08220691
CC GENERAL INFORMATION:
CC APPLICANT: Haseltine, William, et. al.
CC TITLE OF INVENTION: Human Genes, Sequences, and
CC TITLE OF INVENTION: Expression Products
CC NUMBER OF SEQUENCES: 12049
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carella, Byrne, Bain, Giffillan,
CC ADDRESSEE: Cecchi, Stewart, & Olstein
CC STREET: 6 Becker Farm Road
CC CITY: Roseland
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
CC COMPUTER: HP Vectra 486/33
CC OPERATING SYSTEM: MSDOS version 5.0

CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/220,691
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Olstein, Elliot
CC REGISTRATION NUMBER: 24,025
CC REFERENCE/DOCKET NUMBER: 325800-71
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 994-1700
CC TELEFAX: (201) 994-1744
CC INFORMATION FOR SEQ ID NO: 5097:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 495 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
SQ Sequence 495 BP; 97 A; 96 C; 127 G; 131 T; 44 other;

Query Match 11.4%; Score 258; DB 37; Length 495;
Best Local Similarity 84.8%; Pred. No. 4.72e-187;
Matches 352; Conservative 0; Mismatches 52; Indels 11; Gaps 11;

Db 14 GNGCAGTGNAGTTGGCACAGCCACCGCTCTGTGGTGGCTTGTCCCTTAGTCCCGAG 73
QY 28 GCGCAGTG-AGGTTGGCACAGCCACCGCTCTGTGGTGGCTTGTCCCTTAGTCCCGAG 86
Db 74 CGCTCGCCACTGCAGATTCCTTCCCGTGCAGACATGGCCTCTGGCACACACACCG 133
QY 87 CGCTCGCCACTGCAGATTCCTTCCCGTGCAGACATGGCCTCTGGCACACACACCG 146
Db 134 CCGTGTAGATTGNNATATTGTGTGAACAGGCGCTGATGATCCAGAA-TTTAGAAG 192
QY 147 CCGTG-AAGATTGG-AATTAATTGTGGAACAGGCGCTGATGATCCAGAA-TTTAGAAG 204
Db 193 AAGAACTGCCAATATGTGATCTCATTTTGGCAAGCCATCTGATGCTTAATTTGGG 252
QY 205 AAGAACTGAAAAATATGTGATCTCATTTTGGCAAGCCATCTGATGCTTAATTTGGG 264
Db 253 GGAAGNTAAAAATGTTGATTCGCTTCTTGGCAAGCCATGTTAGAGNACACCAT 312
QY 265 G-AAGATAAAAAATGTGATTCATCTCTCTG-CAA-GGCATGGAAGGACACACCAT 321
Db 313 CANGGCGCTTAAAGTTCAATTACAGGGGANNTTTGGNTTTNNAAGAA-AGGTTTN 371
QY 322 CATG-CCTTCAAGGT-CAACTACAGGCGAACATCTGGGCTTTGAAGGAAGAGGCTGT 379
Db 372 NAACCTTTTNGNAGCCNNAANTTNNNGTNNTTGGGGGGGGGNTTTNACC 426
QY 380 ACACATGTCTAGTACCACAGCTT-GTGGCTCCTTGAGGGAGAGATTCAAGCC 433

RESULT 9
ID US-08-196-363-5749 STANDARD; DNA; UNC; 205 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 5749, Application US/08196363.
CC Sequence 5749, Application US/08196363
CC GENERAL INFORMATION:
CC APPLICANT: Kerlavage, Anthony
CC APPLICANT: Fraser, Claire M.
CC APPLICANT: Kirkness, Ewen
CC APPLICANT: Sutton, Granger
CC APPLICANT: Gocayne, Jeanine
CC APPLICANT: Liu, Li-Ing
CC APPLICANT: Fitzgerald, Lisa
CC APPLICANT: Adams, Mark
CC APPLICANT: Lee, Normal
CC APPLICANT: Fuldner, Rebecca

CC APPLICANT: Fleischmann, Robert
CC APPLICANT: Bult, Carol
CC APPLICANT: Blake, Judy
CC APPLICANT: White, Owen
CC APPLICANT: Clayton, Rebecca
CC APPLICANT: Pellegrino, Susan
CC TITLE OF INVENTION: Human Genes, Sequences, and Expression
CC NUMBER OF SEQUENCES: 16818
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Carella, Byrne, Baln, Giffillan, Cecchi, Stewart, &
CC ADDRESSEE: Olstein
CC STREET: 6 Becker Farm Road
CC CITY: Roseland
CC STATE: New Jersey
CC COUNTRY: USA
CC ZIP: 07068
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
CC COMPUTER: HP Vectra 486/33
CC OPERATING SYSTEM: MSDOS version 5.0
CC SOFTWARE: ASCII Text
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/196,363
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Olstein, Elliot
CC REGISTRATION NUMBER: 24,025
CC REFERENCE/DOCKET NUMBER: 325800-91
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (201) 994-1700
CC TELEFAX: (201) 994-1744
CC INFORMATION FOR SEQ ID NO: 5749:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 205 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
SQ Sequence 205 BP; 72 A; 26 C; 31 G; 70 T; 6 other;

Query Match 8.2%; Score 185; DB 34; Length 205;
Best Local Similarity 94.7%; Pred. No. 6.39e-125;
Matches 195; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

Db 1 ATGAACANTCTGATTCCTTGCATTTGGAGGTAATATGTTTATGTTATAT-TAGA 59
QY 1546 ATGAACAATCTGAATCCCTTGCATTTGGAGGTAATATGTTTATATCTAGA 1605
Db 60 TTGTGATCANCAACTGAAATCCCTTTNNCATATTTTCANTGTCCCTAAAGACAGCGT 119
QY 1606 TTGTGAATCAGCAACTGAAATCCCTTTTGCATATTTCAATGTCTTAAAGACAGCGT 1665
Db 120 TGCNCTATATATGAGGTGAAGAAAGATATGTTAGCATTTTATAGTACTAGTTTCTTT 179
QY 1666 TGCTCTATATATGAAGTGAAAAAGATATGTTAGCATTTTATAGTACTAGTTTCTTT 1725
Db 180 AAATGCTATGTAATATACAAAAA 205
QY 1726 AAATGCTATGTAATATACAAAAA 1751

RESULT 10
ID US-08-404-891-322 STANDARD; DNA; UNC; 137 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 322, Application US/08404891.
CC Sequence 322, Application US/08404891
CC GENERAL INFORMATION:
CC APPLICANT: Stuart, Susan G.


```
CC APPLICANT: Delegeane, Angelo M.
CC APPLICANT: Kelleler, Colleen
CC APPLICANT: Selhamer, Jeffrey J.
CC APPLICANT: Wilde, Craig G.
CC APPLICANT: Freitas, Nicole
CC APPLICANT: Bills, Pamela Kay
CC APPLICANT: Pham, Mino Thu
CC APPLICANT: Altus, Christina M.
CC TITLE OF INVENTION: PLACENTA DERIVED POLYNUCLEOTIDES AND POLYPEPTIDES
CC NUMBER OF SEQUENCES: 2870
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3330 HILLVIEW AVENUE
CC CITY: PALO ALTO
CC STATE: CALIFORNIA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/404,891
CC FILING DATE: Herewith
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: LUTHER, BARBARA J.
CC REGISTRATION NUMBER: 33954
CC REFERENCE/DOCKET NUMBER: PD 0030 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 855-0555
CC TELEFAX: (415) 855-0572
CC INFORMATION FOR SEQ ID NO: 322:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 137 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: 179320
SQ Sequence 137 BP; 24 A; 41 C; 36 G; 30 T; 6 other;

Query Match 5.2%; Score 119; DB 55; Length 137;
Best Local Similarity 94.9%; Pred. No. 4.84e-70;
Matches 131; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Db 1 GCCAGTGAGGTGGCAGACG-ACNGCTCTGTGGCTCGCTTGGTTCCCTTAGTCCCGAGC 59
   |||||||
Qy 28 GCGCAGTGAGGTGGCAGACGCCACCGCTCTGTGGCTCGCTTGGTTCCCTTAGTCCCGAGC 87
      60 GCTGNNCACATGCAGATTCTTCCTCCCGTGACACATGGCCTCTGG-AACCACCACCGN 118
         |||||
Qy 88 GCTGCCCACTGCAGATTCCTTCCCGTGACACATGGCCTCTGGCACCAACCACCGCC 147
      119 CGTGAAGATTGGATAAT 136
          |||||||
Qy 148 CGTGAAGATTGGATAAT 165

RESULT 11
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC TITLE OF INVENTION: 5-PHOSPHATASE
```

```

CC      NUMBER OF SEQUENCES: 38
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: CHIRON CORPORATION
CC      STREET: 4560 HORTON STREET
CC      CITY: EMERYVILLE
CC      STATE: CA
CC      COUNTRY: USA
CC      ZIP: 94608
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/624,190
CC      FILING DATE:
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: GREEN, GRANT
CC      REGISTRATION NUMBER: 31,259
CC      REFERENCE/DOCKET NUMBER: 1182.003
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 510-601-2706
CC      TELEFAX: 510-655-3542
CC      INFORMATION FOR SEQ ID NO: 14:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 578 base pairs
CC      TYPE: nucleic acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

```

[illegible]

Db 539 RKSRSNTTRVKSAVHSHSKGRDYRD 563
OY 886 ACCTCAGATAGGGTCCACAGAAATGG 910

RESULT 12
ID US-08-624-190-14 STANDARD; DNA; UNC; 578 BP.

AC xxxxxx
DT 01-JAN-1900
DE Sequence 14, Application US/08624190.
CC Sequence 14, Application US/08624190.
CC GENERAL INFORMATION:
CC APPLICANT: KAVANAUGH PH.D., MICHAEL
CC APPLICANT: POT PH.D., DAVID
CC APPLICANT: WILLIAMS PH.D., LEWIS T.
CC TITLE OF INVENTION: SIGNALING INOSITOL POLYPHOSPHATE
CC TITLE OF INVENTION: 5-PHOSPHATASE
CC NUMBER OF SEQUENCES: 38
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: CHIRON CORPORATION
CC STREET: 4560 HORTON STREET
CC CITY: EMERYVILLE
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94608
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/624,190
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: GREEN, GRANT
CC REGISTRATION NUMBER: 31,259
CC REFERENCE/DOCKET NUMBER: 1182.003
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 510-601-2706
CC TELEFAX: 510-655-3542
CC INFORMATION FOR SEQ ID NO: 14:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 578 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC Sequence 578 BP; 37 A; 11 C; 62 G; 65 T; 403 other;

Query Match 3.4%; Score 78; DB 66; Length 578;
Best Local Similarity 13.0%; Pred. No. 2.59e-37;
Matches 56; Conservative 195; Mismatches 177; Indels 3; Gaps 3;

Db 31 DKVKAHSHRSVTVKASGKNKVDVSGKKKSDGSDKYSHKKSKNKVVTKKRYVADSCK 90
Cp 1080 GGGCATGAAGCTGCATGCTCTTCCCATATTTTTCAGTTAAGCAGGAATCCCAAAA 1021
Db 91 RGCMMKNHSDMTGTWNMGNAKKTWSKGGKTRDDSDADYHDYVETDSKWKHSTVTKTVAH 150
Cp 1020 TGACTGGAAATTGAATGCTCTTCTTCTCCTGCGGACGACCATGCTTAATGCTTGGT 961
Db 151 TWMRYVAKHNRSHCTDNVNTGANTGNKGAAGVSMNGTSGVNSHTSGSKRRNNYMRAGD 210
Cp 960 AATAAACAAGAAACTGGGCCATATTCCTTCAGGTATGAGAGGTTTCTGACCATTCCTGTG 901
Db 211 KKSNTHRTHWGDNRYVDVNTATKKYADSHDTRKRVHTATYRTRD-KYAYTKKATGMKYNS 269
Cp 900 GACCCATCTGAGTATGAGTAGAGCAGTAACTTTTGGCTTATATAGCGTTTCTTTC 841
Db 270 WCDRVKSVYHVVCYGSYSDMTSDHSVATAGVTSVSKNGTVDSGRCYATKTKSTKYHS 329
Cp 840 AGGGTCTTTAAGACCCCGGTCCACCGAAACCTGCTTCCGTGCTCCTTCCAGCAGTCATATA 781

Db 330 SCSVKSNGSGVVKTKKSDYDHSKSSDSYGGCARATYTHHGTGHGKTSKTRKYD 389
Cp 780 TCTGTCCCATGCGCATCTGCGTATACAAATTCAGCCTCCTTAGCAAGAACCACCTCT 721
Db 390 VKTR-DSSGKTKSTSHDMKWTSRACSGSSSTNNYMGVGMHYKTSDF-AWSYDKDSCRG 447
Cp 720 GGAAGTGTCATGTGTGATACATCCGCCGCCAGGTGCGGACATGAAGCTTCTGCC 661
Db 448 STGSKKSTANR 458
Cp 660 CGGAGCTAAA 650

RESULT 13
ID US-07-904-074A-10 STANDARD; DNA; UNC; 477 BP.

AC xxxxxx
DT 01-JAN-1900
DE Sequence 10, Application US/07904074A.
CC Sequence 10, Application US/07904074A.
CC GENERAL INFORMATION:
CC APPLICANT: John E. Shively
CC APPLICANT: Rainer Fischer
CC APPLICANT: Anna Wu
CC APPLICANT: Roy Paxton
CC APPLICANT: Y.H. Joy Yang
CC TITLE OF INVENTION: Chimeric Anti-CEA Antibody
CC NUMBER OF SEQUENCES: 10
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: City of Hope
CC STREET: 1500 East Duarte Road
CC CITY: Duarte
CC STATE: California
CC COUNTRY: United States of America
CC ZIP: 91010-0269
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: 3M Double Density 5 1/4" diskette
CC COMPUTER: Wang PC
CC OPERATING SYSTEM: MS-DOS (R) Version 3.30
CC SOFTWARE: Microsoft (R)
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/904,074A
CC FILING DATE: 19920615
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA: None
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Irons, Edward S.
CC REGISTRATION NUMBER: 16,541
CC REFERENCE/DOCKET NUMBER: None
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 785-6938
CC TELEFAX: (202) 785-5351
CC TELEX: 440087 LM WSH
CC INFORMATION FOR SEQ ID NO: 10:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 477
CC TYPE: NUCLEIC ACID
CC STRANDEDNESS: Single Stranded
CC TOPOLOGY: Unknown
CC MOLECULE TYPE: Nucleic Acid
CC HYPOTHETICAL: Not Applicable
CC ANTI-SENSE: Not Applicable
CC FRAGMENT TYPE: Not Applicable
CC ORIGINAL SOURCE: Synthetically Prepared
CC IMMEDIATE SOURCE: Synthetically Prepared
CC POSITION IN GENOME: None
CC FEATURE: None
CC PUBLICATION INFORMATION: None
CC Sequence 477 BP; 23 A; 12 C; 35 G; 35 T; 372 other;

Query Match 3.2%; Score 73; DB 13; Length 477;
Best Local Similarity 8.5%; Pred. No. 1.86e-33;
Matches 39; Conservative 154; Mismatches 263; Indels 1; Gaps 1;

Db	5	HRHNTMNNNGSNNNNNVNVNKGVNCNRYKNVNSGGGNVKNNGSGSNKNSCSAASGNTNSSIAMS	64
QY	21	CACTCCCGCGCAGTGAAGTTGGCACAGACCAGCTCTGTGGCTCGCTTGGTTCCCTTAGT	80
Db	65	VRNTNNKRNNVVAASNSSDGTNTYVDSVYGRNTVSRDNARNNNYNNMSSNRSNDTAMYCA	124
QY	81	CCCCAGCGCTCGCCCACTGCAGATTCTTTCCCGTGACAGATGGCCTCTGGCACCACA	140
Db	125	RNDYGGGGNGYNGT NATVSAASTKGSVNNNNANSSKSTSGGTAANGCNV KDYNNNV	184
QY	141	CCACCGCGGTGAAGATTGGAATAATTGGTGAACAGGCTGTGATGATCCAGAAATTTAG	200
Db	185	TVSWNSGANTSGVHTNNAVNVNSSGNYSNSVTVTVNSSNGTNTYCNVNHKNSNTKYDK	244
QY	201	AAGAGAAGAACTGAAAAATATGTGATACATCCATTGTGGCAAGCACTGTGATGCCTTAATT	260
Db	245	VNNKSCDKHTCCNNCANNNNGNSVNNNNNNKNDTNNNSRTNNYTCVVVDASHNDNVK	304
QY	261	TGGGGAAGATAAAAAATGTGATTGCATCTCTCTTGC AAGGCATGGAAGCACACACA	320
Db	305	NNMYVDGVNVHNAKTKNRNNNYNSTYRVVSVTVNHNDWNGKNYKCKVSNKANANNNK	364
QY	321	TCATGCCCTCAAGGTCAAC-TACCAGGCGAACATCTGGGCTTTGAAGGAAGAGGGCTGT	379
Db	365	TNSKAKGNRRNNYYTTNNNSRDNTKNNVNTCNVKGNTNSDNAVWNNSNGNNNNYTT	424
QY	380	ACACATGTCAATGACACACACAGCTGTGGCTCTTGAGGGAGGAGATTACAGCCCGCGAT	439
Db	425	NNVNSDGSNNNYSKNTVDKSRMNGNVN SCSVMHNA	461
QY	440	ATTGCAATTAATGATCAGTTCATTCATGACAGACCACTA	476

RESULT 14
ID US-08-446-939-8 STANDARD; DNA; UNC; 383 BP.
AC xxxxxx
DT 01-JAN-1900
DE Sequence 8, Application US/08446939.
CC Sequence 8, Application US/08446939
CC GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Millitia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentln Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,939
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC

CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WHI91-09FX
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
CC Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match	2.9%;	Score 66;	DB 59;	Length 383;
Best Local Similarity	11.9%;	Pred. No. 3.95e-28;		
Matches	43;	Conservative	165;	Mismatches 149;
			Indels	4;
			Gaps	4;

```

Db      1 MTSSGVCVCHGSGAAGWHVYTRASTHYKSYNNDMVTDNNNGAAYKCKDYRSTCDNKSCHSANC 60
        :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      697 ATCCGCCCCCAGGTGCGGAACATGAAGCTTTCTGCCCCGGAGCTAAACGAGGTCCCTC 638
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      61 TSCKVCAVVRKRDNTTVCHDKYHDDAASKCMKKKGTMCSGSSDCNDNSYNTSNDVVTG 120
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      637 GATTGTGACCATGTGTCCCTTTGAGTGGCACCGGAGTCTTAAGCTTCTTAGCAGTCTC-TA 579
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      121 SGVASVTCYRVNRKSTWTGKTRKMSHCADDRSDSSTCANNNHNTDTVGKRAVYRAKN 180
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      578 T-AAGAACCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATGGAATA-TGGCACACTCC 521
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      181 TSTVAVKYDHYASWKRKDSNDKNHNTARKTGKYWTAHAAGNYTRHVSWDRNVGSSARGSH 240
        |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      520 TCTGGCACAAGATGACTTCCATCATAGAAGACTGAGGTCTCATAGTGTGCTCTGCAAT 461
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      241 HSDHTCGRKMYHRDKSSNVKNDTCCCDGSRGYSSVDANGSYGTARYMAVSRMNNASKTD 300
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      460 GAACGTGATCAATAATGACATATCGCGGGCTGATCTCTCCCT-CAAGAGGCCACAAG 402
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      301 VYSMAVWMTSRCAVGVKDYGSKVRDYYSMKDNVRDRGRNSWNHGMVCTTCWDHDART 360
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Cp      401 CTGTGGTCACTATGACATGTGTACAGCCCTCTTCCTTCAAGCCCAAGATGTTCCGCTGGT 342
        :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      361 A 361
        |
Cp      341 A 341

```

```

CC ID      US-08-446-936-8 STANDARD; DNA; UNC; 383 BP.
CC AC      xxxxxx
CC DT      01-JAN-1900
CC DE      Sequence 8, Application US/08446936.
CC CC      Sequence 8, Application US/08446936
CC          GENERAL INFORMATION:
CC APPLICANT: Lin, Herbert Y.
CC APPLICANT: Wang, Xiao-Fan
CC APPLICANT: Weinberg, Robert A.
CC APPLICANT: Lodish, Harvey F.
CC TITLE OF INVENTION: TGF- Type Receptor cDNAs Encoded
CC TITLE OF INVENTION: Products and Uses Therefor
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
CC STREET: Two Militia Drive
CC CITY: Lexington
CC STATE: Massachusetts
CC COUNTRY: U.S.A.
CC ZIP: 02173
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

```

CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/446,936
CC FILING DATE:
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/311,703
CC FILING DATE: 23-SEP-1994
CC CLASSIFICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/786,063
CC FILING DATE: 31-OCT-1991
CC CLASSIFICATION: 530
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Granahan, Patricia
CC REGISTRATION NUMBER: 32,227
CC REFERENCE/DOCKET NUMBER: WH191-09FY
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-861-6240
CC TELEFAX: 617-861-9540
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 383 base pairs
CC TYPE: nucleic acid
CC STRANDEDNESS: double
CC TOPOLOGY: linear
CC MOLECULE TYPE: DNA (genomic)
SQ Sequence 383 BP; 27 A; 25 C; 26 G; 34 T; 271 other;

Query Match 2.9%; Score 66; DB 59; Length 383;
Best Local Similarity 11.9%; Pred. No. 3.95e-28;
Matches 43; Conservative 165; Mismatches 149; Indels 4; Gaps 4;
Db 1 MTSSGVCHSGAAGMHWTRASTHVKSVNNNDVTDNNGAVKCKDVRSTCDNKSMSNCS 60
CP 697 ATCCGCCGCCAGGTGCGGAACATGAAGCTTCTGCCCGGAGCTAAACGAGGTCCCTC 638
Db 61 TSCKVCYAVWRKNDNTVCHDKYHDDAASKMKKKKGTMCSSDCNDNSYNTSNDVVTG 120
CP 637 GATTGTGACCATGTGCCCTTTGAGTGCGACCGGAGTCCCTAGCTCTTAGCAGTCTC-TA 579
Db 121 SGVASYCYRVNRKSSWTGKTRKMSHCADRSDSSTCANNHNHTDVGKRAYKKAKN 180
CP 578 T-AAGAACCTCTCTCGTTTGGGGCAAAACGGCTCAGCCATTGGAATA-TGGCACACTCC 521
Db 181 TSTVAVKYDYHYSWKDRKSDNKHNTARKTKKYWTAHAKGNYTRHVSMDRNVGSSARSH 240
CP 520 TCTGGACAAGAATGACTTCATCATAGAGAGACTGAGGTCTCATAGTGTCTGTCAT 461
Db 241 HSDHTCGRKMVHRDKSSNVKNDTCCCDGSRGYSVDDANSVGTARYMAVSRMNNASKTD 300
CP 460 GAACGTGATCAATATGACATATATCGCGGGCTGAATCTCTCCCT-CAAGGAGCCACAAG 402
Db 301 VYSMAVMTSRCNAVGVKDYGSKVRDVSMKDNVRDRGTRNSSWNHGMVCTTCWDHART 360
CP 401 CTGTGTCACATATGACATGTGTACAGCCCTCTTCTTCAAGCCCAAGATGTTGCGCTGT 342
Db 361 A 361
CP 341 A 341

Search completed: Mon Feb 10 23:43:38 1997
Job time : 1639 secs.